ORIGINAL ARTICLE



# CrossMark

# **Environmental Induction of Neurodevelopmental** Disorders

Rodrick Wallace<sup>1</sup>

Received: 7 July 2016 / Accepted: 14 October 2016 / Published online: 26 October 2016 © Society for Mathematical Biology 2016

Abstract Chemical exposures, pre- and neonatal infections, psychosocial stress, and the cross-generational cultural and epigenetic impacts of these and other toxicants become an integrated, sometimes synergistic, signal that can overwhelm essential neurodevelopmental regulation. We characterize that dynamic through statistical models based on the asymptotic limit theorems of control and information theories. Schizophrenia and autism emerge as two different 'phases' of pathological neurodevelopmental 'condensations' that impair the dynamic, shifting global workspace of normal mental function.

Keywords Autism · Control theory · Data rate theorem · Information theory · Mitochondrial free energy · Phase transition · Rate-distortion theorem · Schizophrenia

Early neglect and life course environmental insults ... can lead to impaired neuronal responsiveness and symptoms of profound prefrontal cortical dysfunction, providing a direct link between the environment and the cognitive impairments observed in psychotic syndrome. (Os et al. 2010)

## **1** Introduction

Growing horror over the possible relation between the rapid spread of mosquitoborne Zika virus and increased rates of microcephaly among children born to women infected during pregnancy (Hayden 2016; Garcez et al. 2016) sharply focuses attention on how environmental exposures, in a large sense, might trigger neurodevelopmental

Rodrick Wallace Wallace@nyspi.columbia.edu; rodrick.wallace@gmail.com

Division of Epidemiology, The New York State Psychiatric Institute, New York, NY, USA

disorders. Here, we extend the perspective of Wallace (2015a), which focused on the pathologies of aging, to examine the induction of developmental disorders in the presence of environmental disruption, in a large sense.

The underlying importance of neurodevelopment in the etiology of serious mental disorder has been described cogently by Corbin et al. (2008) who conclude that

... [U]nraveling the mechanisms of neural progenitor cell diversity in the brain has tremendous clinical importance ... [D]efects in any of these processes can have devastating and long lasting consequences on brain function ...

... [A]bnormal development of interneurons may be an underlying causative factor, or contribute to the phenotype of a variety of developmental disorders ... [including] autism spectrum disorders ... [and] schizophrenia ...

Similarly, Tiberi et al. (2012) describe how the cerebral cortex is composed of hundreds of different types of neurons, which underlie its ability to perform highly complex neural processes. How cortical neurons are generated during development constitutes a major challenge in developmental neurosciences with important implications for brain repair and diseases. Cortical neurogenesis is dependent on intrinsic and extrinsic clues, which interplay to generate cortical neurons at the right number, time, and place. Recent evidence, in their view, indicates that most classical morphogens, produced by various neural and nonneural sources throughout embryonic development, contribute to the master control and fine-tuning of cortical neurogenesis. They conclude that the molecular control of cortical neurogenes involves the interplay of intrinsic and extrinsic cues that coordinate the pattern of neural progenitor division and differentiation.

Rapoport et al. (2012) find the neurodevelopmental model positing illness as the end stage of abnormal neurodevelopmental processes that began years before the onset of the illness. Environmental risk factors such as urbanicity, childhood trauma, and social adversity have received strong replication with marked phenotypic nonspecificity pointing to common brain development pathways across disorders. The neurodevelopmental model of schizophrenia has long existed as a model for other childhood-onset conditions, including attention-deficit hyperactivity disorder, intellectual deficiency, autism spectrum disorders (ASD), and epilepsy.

Rapoport et al. specifically identify infection/famine, placental pathology, low birth weight, urban environment, childhood trauma, and ethnic minority/immigrant status in disease etiology, concluding that in the central nervous system, neuronal proliferation, cell migration, morphological and biochemical differentiation, and circuit formation all depend on cell and cell–environment interactions that control developmental process, and so can cause altered trajectories. Their Fig. 1 provides a summary schematic.

A parallel line of argument explores mitochondrial abnormalities that are closely associated with both schizophreniform and autism spectrum disorders.

Ben-Shachar (2002) finds mitochondrial impairment could provide an explanation for the broad spectrum of clinical and pathological manifestations in schizophrenia. Several independent lines of evidence, Ben-Shachar asserts, suggest an involvement of mitochondrial dysfunction in the disorder, including altered cerebral energy metabolism, mitochondrial polyplasia, dysfunction of the oxidative phosphorylation system, and altered mitochondrial-related gene expression. They conclude that the interaction between dopamine, a predominant etiological factor in schizophrenia, and mitochondrial respiration is a possible mechanism underlying the hyper- and hypoactivity cycling in schizophrenia.

Prabakaran (2004) claim that almost half the altered proteins identified by a brain tissue protenomics analysis of samples from schizophrenic patients were associated with mitochondrial function and oxidative stress responses. They propose that oxidative stress and the ensuing cellular adaptations are linked to the schizophrenia disease process.

Shao (2008) similarly find evidence of mitochondrial dysfunction in schizophrenia. Likewise, Scaglia (2010) suggests involvement of mitochondrial dysfunction in schizophrenia and argues that mechanisms of dysfunctional cellular energy metabolism underlie the pathophysiology of major subsets of psychiatric disorders.

Clay et al. (2011) point to an underlying dysfunction of mitochondria in bipolar disorder and schizophrenia including (1) decreased mitochondrial respiration; (2) changes in mitochondrial morphology; (3) increases in mitochondrial DNA (mtDNA) polymorphisms and in levels of mtDNA mutations; (4) downregulation of nuclear mRNA molecules and proteins involved in mitochondrial respiration; (5) decreased high-energy phosphates and decreased pH in the brain; and (6) psychotic and affective symptoms, and cognitive decline in mitochondrial disorders. They conclude that understanding the role of mitochondria, both developmentally and in the ailing brain, is of critical importance to elucidate pathophysiological mechanisms in psychiatric disorders.

There is likewise considerable and growing evidence for mitochondrial mechanisms in autism spectrum disorders (ASD).

Palmieri and Persico (2010) find ASD often associated with clinical, biochemical, or neuropathological evidence of altered mitochondrial function. The majority of autistic patients displays functional abnormalities in mitochondrial metabolism seemingly secondary to pathophysiological triggers. Thus, in their view, mitochondrial function may play a critical role not just in rarely causing the disease, but also in frequently determining to what extent different prenatal triggers will derange neurodevelopment and yield abnormal postnatal behavior.

Giulivi (2010) similarly assert that impaired mitochondrial function may influence processes highly dependent on energy, such as neurodevelopment, and contribute to autism. In their study, children with autism were more likely to have mitochondrial dysfunction, mtDNA overreplication, and mtDNA deletions than typically developing children.

A long series of studies by Rossignol and Frye (2010) find evidence accumulating that autism spectrum disorder is characterized by certain physiological abnormalities, including oxidative stress, mitochondrial dysfunction, and immune dysregulation/inflammation. Recent studies, they conclude, have reported these abnormalities in brain tissue derived from individuals diagnosed with ASD as compared to brain tissue derived from control individuals, suggesting that ASD has a clear biological basis with features of known medical disorders.

Goh et al. (2014) argue that impaired mitochondrial function impacts many biological processes that depend heavily on energy and metabolism and can lead to a wide range of neurodevelopmental disorders, including autism spectrum disorder. Although, in their view, evidence that mitochondrial dysfunction is a biological subtype of ASD has grown in recent years, no study had previously demonstrated evidence of mitochondrial dysfunction in brain tissue in vivo in a large, well-defined sample of individuals with ASD. Their use of sensitive imaging technologies allowed them to identify in vivo a biological subtype of ASD with mitochondrial dysfunction. Lactatepositive voxels in their sample were detected most frequently in the cingulate gyrus, a structure that supports higher-order control of thought, emotion, and behavior, and one in which both anatomical and functional disturbances have been reported previously in ASD.

For neurodevelopment, control of gene expression is everything, and mechanisms by which environmental factors interfere with control are of essential clinical and epidemiological concern.

Here, we will describe statistical models of developmental failure based on the asymptotic limit theorems of control and information theories that may provide new tools in exploring such mechanisms. The models are analogous to more familiar empirical least-squares regression and may permit deep scientific inference arising from comparison of similar systems under different, or different systems under similar, experimental or observational circumstances.

## 2 A Control Theory Model

It is well understood that there is no gene expression without regulation. This implies that gene expression is inherently unstable in the formal control theory sense of the data rate theorem (Nair et al. 2007) and must be stabilized by provision of control information at a critical rate. Failure to provide control information at or above that rate initiates characteristic modes of system failure that, for neural systems, are expressed as developmental disorders. More explicitly, assuming an approximate nonequilibrium steady state, the simplest 'regression' model of deviations from that state—described in terms of an *n*-dimensional vector of observables  $x_t$  at time *t*—has the form

$$x_{t+1} = \mathbf{A}x_t + \mathbf{B}u_t + W_t \tag{1}$$

where  $x_{t+1}$  is the state at time t + 1,  $u_t$  is the imposed *n*-dimensional control signal vector at time t,  $W_t$  is an added noise signal, and **A** and **B** are, in this approximation, fixed  $n \times n$  matrices. See Fig. 1 for a schematic.

The data rate theorem (Nair et al. 2007) states that, for an inherently unstable system, the control information represented by the vector  $u_t$  must be provided at a rate  $\mathcal{H}$  that is greater than the rate at which the system produces 'topological information.' For the system of Eq. (1) and Fig. 1, that rate is given as

$$\mathcal{H} > \log[|\det(\mathbf{A}^{u})|] \equiv \alpha_0 \tag{2}$$

where det is the determinant and  $\mathbf{A}^{u}$  is the component submatrix of  $\mathbf{A}$  that has eigenvalues  $\geq 1$ .



**Fig. 1** 'Regression model' for a control system near a nonequilibrium steady state.  $x_t$  is system output at time t,  $u_t$  the control signal, and  $W_t$  an added noise term

An alternate derivation of Eq. (2) is given in Sect. 6.

Generalization to more complex inherently unstable systems in the context of a scalar integrated environmental insult  $\rho$ —for example, taken as the magnitude of the largest vector of an empirical principal component analysis—suggests that Eq. (2) will become something like

$$\mathcal{H}(\rho) > f(\rho)\alpha_0 \tag{3}$$

 $f(0)\alpha_0$  is then interpreted as the rate at which the system generates topological information in the absence of an integrated environmental exposure.

What are the forms of  $\mathcal{H}(\rho)$  and  $f(\rho)$ ? In Sect. 6, we calculate  $\mathcal{H}(\rho)$  as the 'cost' of control information, given the 'investment'  $\rho$ , using a classic Black–Scholes approximation (Black and Scholes 1973). To first order,

$$\mathcal{H}(\rho) = \kappa_1 \rho + \kappa_2 \tag{4}$$

where the  $\kappa_i$  are positive or zero.

If we take the same level of approximation,  $f(\rho)$  in Eq. (3) can be similarly expressed as  $\kappa_3\rho + \kappa_4$  so that the stability condition is

$$\mathcal{T} \equiv \frac{\kappa_1 \rho + \kappa_2}{\kappa_3 \rho + \kappa_4} > \alpha_0 \tag{5}$$

For small  $\rho$ , the stability requirement is  $\kappa_2/\kappa_4 > \alpha_0$ , and at high  $\rho$  it becomes  $\kappa_1/\kappa_3 > \alpha_0$ . If  $\kappa_2/\kappa_4 \gg \kappa_1/\kappa_3$ , then at some intermediate value of  $\rho$ , the essential inequality may be violated, leading to failure of neurodevelopmental regulation. See Fig. 2.

Fetal, child, and indeed adult developmental trajectories are embedded not only in environments of direct exposure, but of multimodal inheritance, both through



cross-generational gene methylation and other biochemical mechanisms and via sociocultural influences. It is, however, implicit that direct environmental exposures, inherited gene methylation, sociocultural inheritance, and other important factors must interact along the developmental trajectory. Thus, rather than a simple scalar, we are confronted by an  $m \times m$  matrix having elements  $\rho_{i,j}$  i,  $j = 1 \dots m$ .

Square matrices of order *m*, however, have *m* scalar invariants, *m* real numbers that characterize the matrix regardless of how it is expressed in different coordinate systems. The first is the trace, and the last  $\pm$  the determinant. In general, the invariants are the coefficients of the characteristic polynomial  $\mathcal{P}(\lambda)$ :

$$\mathcal{P}(\lambda) = \det(\rho - \lambda I)$$
  
=  $\lambda^m + r_1 \lambda^{m-1} + \dots + r_{m-1} \lambda + r_m$  (6)

where  $\lambda$  is a parameter that is an element of some ring, det is the determinant, and *I* the *m* × *m* identity matrix. Note that  $\lambda$  may, in fact, be taken as the matrix  $\rho$  itself, since square matrices form a ring, in which case the relation is a matrix polynomial  $\mathcal{P}(\rho) = 0 \times I$ .

For a  $m \times m$  matrix, we have invariants  $r_1, \ldots, r_m$  and an appropriate scalar ' $\rho$ ' in Eq. (6)—determining the 'temperature'  $\mathcal{T}$ —is then a monotonic increasing function of the  $r_i$ :

$$\hat{\rho} = \hat{\rho}(r_1, \dots, r_m) \tag{7}$$

so that

$$\mathcal{T}(\hat{\rho}) = \frac{\kappa_1 \hat{\rho} + \kappa_2}{\kappa_3 \hat{\rho} + \kappa_4} \tag{8}$$

🖉 Springer

We have invoked the 'rate-distortion manifold' of Glazebrook and Wallace (2009)—formulated as a 'generalized retina' in Wallace and Wallace (2010)—to project a complicated 'information manifold' down onto a lower dimensional 'tangent space' tuned across that manifold in such a way as to preserve most of the underlying information. Here, we assume a scalar tangent space. Higher dimensional structures are possible in a standard manner at the cost of some considerable increase in mathematical overhead.

What are the dynamics of  $\mathcal{T}(\hat{\rho})$  under stochastic circumstances? We explore this by examining how a control signal  $u_t$  in Fig. 1 is expressed in the system response  $x_{t+1}$ . More explicitly, we suppose it possible to deterministically retranslate a sequence of system outputs  $X^i = x_1^i, x_2^i, \ldots$  into a sequence of possible control signals  $\hat{U}^i = \hat{u}_0^i, \hat{u}_1^i, \ldots$  and then compare that sequence with the original control sequence  $U^i = u_0^i, u_1^i, \ldots$  The difference between them is a real number measured by a chosen distortion measure, enabling definition of an average distortion

$$\langle d \rangle = \sum_{i} p(U^{i})d(U^{i}, \hat{U}^{i})$$
(9)

where (1),  $p(U^i)$  is the probability of the sequence  $U^i$ , (2)  $d(U^i, \hat{U}^i)$  is the distortion between  $U^i$  and  $\hat{U}^i$ , and (3) the sequence of control signals has been deterministically reconstructed from the system output.

It then becomes possible to apply a classic rate–distortion theorem (RDT) argument. According to the RDT, there exists a rate–distortion function (RDF) that determines the minimum channel capacity, R(D), necessary to keep the average distortion <d>below some fixed limit D (Cover and Thomas 2006). Based on Feynman (2000) interpretation of information as a form of (free) energy, we can then construct a Boltzmann-like pseudoprobability in the 'temperature' T as

$$dP(R, \mathcal{T}) = \frac{\exp[-R/\mathcal{T}]dR}{\int_0^\infty \exp[-R/\mathcal{T}]dR}$$
(10)

since higher T necessarily implies greater channel capacity.

The integral in the denominator is essentially a statistical mechanical partition function, and we can then define a 'free energy' Morse function F (Pettini 2007) as

$$\exp[-F/\mathcal{T}] = \int_0^\infty \exp[-R/\mathcal{T}] \mathrm{d}R = \mathcal{T}$$
(11)

so that  $F(\mathcal{T}) = -\mathcal{T} \log[\mathcal{T}]$ .

Then, an entropy analog can also be defined as the Legendre transform of F:

$$S \equiv F(T) - T dF/dT = T$$
<sup>(12)</sup>

As a first approximation, Onsager's treatment of nonequilibrium thermodynamics (Groot and Mazur 1984) can be applied, so that system dynamics are driven by the

gradient of S in essential parameters—here T—under conditions of noise. This gives a stochastic differential equation

$$d\mathcal{T}_t \approx (\mu d\mathcal{S}/d\mathcal{T}) dt + \beta \mathcal{T}_t dW_t = \mu dt + \beta \mathcal{T}_t dW_t$$
(13)

where  $\mu$  is a 'diffusion coefficient' representing the efforts of the underlying control mechanism, and  $\beta$  is the magnitude of an inherent impinging white noise  $dW_t$  in the context of volatility, i.e., noise proportional to signal.

Applying the Ito chain rule to  $\log(\mathcal{T})$  in Eq. (13), a nonequilibrium steady-state (nss) expectation for  $\mathcal{T}$  can be calculated as

$$E(\mathcal{T}_t) \approx \frac{\mu}{\beta^2/2} \tag{14}$$

Again,  $\mu$  is interpreted as indexing the attempt by the embedding control apparatus to impose stability—raise  $\mathcal{T}$ . Thus, impinging noise can significantly increase the probability that  $\mathcal{T}$  falls below the critical limit of Fig. 2, initiating a control failure.

However,  $E(\mathcal{T})$  is an expectation, so that, in this model, there is always some nonzero probability that  $\mathcal{T}$  will fall below the critical value  $\alpha_0$  in the multimodal expression for  $\mathcal{T}(\hat{\rho})$ : sporadic control dysfunctions have not been eliminated. Raising  $\mu$  and lowering  $\beta$  decreases their probability, but will not drive it to zero in this model, a matter of some importance for population rates of neurodevelopmental disorders.

### 3 A 'Cognitive' Model

A different approach to the dynamics of neurodevelopmental regulation applies the 'cognitive paradigm' of Atlan and Cohen (1998), who recognized that the immune response is not merely an automatic reflex, but involves active choice of a particular response to insult from a larger repertoire of possible responses. Choice reduces uncertainty and implies the existence of an underlying information source (Wallace 2012, 2015a, b).

Given an information source associated with an inherently unstable, rapidly acting cognitive neurodevelopmental control system—called 'dual' to it—an equivalence class algebra can be constructed by choosing different system origin states  $a_0$  and defining the equivalence of two subsequent states at times m, n > 0, written as  $a_m$ ,  $a_n$ , by the existence of high-probability meaningful paths connecting them to the same origin point. Disjoint partition by equivalence class, analogous to orbit equivalence classes in dynamical systems, defines a symmetry groupoid associated with the cognitive process. Groupoids are deep generalizations of the group concept in which there is not necessarily a product defined for each possible element pair (Weinstein 1996).

The equivalence classes define a set of cognitive dual information sources available to the inherently unstable neurodevelopment regulation system, creating a large groupoid, with each orbit corresponding to a transitive groupoid whose disjoint union is the full groupoid. Each subgroupoid is associated with its own dual information source, and larger groupoids will have richer dual information sources than smaller. Let  $X_{G_i}$  be the control system's dual information source associated with the groupoid element  $G_i$ , and let Y be the information source associated with embedding 'normal' environmental variation that impinges on development. Wallace (2012, 2015b) gives details of how environmental regularities imply the existence of an environmental information source that, for humans, particularly includes cultural and socioeconomic factors (e.g., Wallace 2015c).

We can again construct a 'free energy' Morse function (Pettini 2007). Let  $H(X_{G_i}, Y) \equiv H_{G_i}$  be the joint uncertainty of the two information sources. Another Boltzmann-like pseudoprobability can then be written as

$$P[H_{G_i}] = \frac{\exp[-H_{G_i}/\mathcal{T}]}{\sum_j \exp[-H_{G_j}/\mathcal{T}]}$$
(15)

 $\mathcal{T}$  is the 'temperature' from Eq. (9), via the  $\hat{\rho}$  of Eq. (7), and the sum is over the different possible cognitive modes of the full system.

A new Morse function  $\mathcal{F}$  is defined by

$$\exp[-\mathcal{F}/\mathcal{T}] \equiv \sum_{j} \exp[-H_{G_{j}}/\mathcal{T}]$$
(16)

Given the inherent groupoid structure as a generalization of the simple symmetry group, it becomes possible to apply an extension of Landau's picture of phase transition (Pettini 2007). In Landau's 'spontaneous symmetry breaking,' phase transitions driven by temperature changes occur as alteration of system symmetry, with higher energies at higher temperatures being more symmetric.

For this model, the shift between symmetries is highly punctuated in the temperature index T under the data rate theorem for unstable control systems. Typically, there are only a very limited number of possible phases, which may or may not coexist under particular circumstances.

Decline in  $\mathcal{T}$  can lead to punctuated decline in the complexity of cognitive process possible within the neurodevelopmental control system, driving it into a ground-state collapse in which neural systems fail to develop normally.

The essential feature is the integrated environmental insult  $\hat{\rho}$ . Most of the topology of the inherently unstable neurodevelopmental system has been 'factored out' so that  $\hat{\rho}(r_1, \ldots, r_m)$  remains the only possible index of the rate of topological information generation for the DRT. Thus, in Eqs. (15) and (16),  $\mathcal{T}(\hat{\rho})$  is again the driving parameter.

Increasing  $\hat{\rho}$  is then equivalent to lowering the 'temperature'  $\mathcal{T}$ , and the system passes from high symmetry 'free flow' to different forms of 'crystalline' structure —broken symmetries representing the punctuated onset of significant neurodevelopmental failure.

Again, if  $\kappa_2/\kappa_4 \gg \kappa_1/\kappa_3$  in Eq. (8), accumulated environmental insult will quickly bring the effective 'temperature' below some critical value, raising the probability for, or triggering the collapse into, a dysfunctional ground state of low symmetry in which essential network connections are not made or else become locally overconnected and globally disjoint. Sufficient conditions for the intractability—stability—of the pathological ground state can be explored using the methods of Wallace (2016). Given a vector of parameters characteristic of and driving that phase, say **J**, that measures deviations from a nonequilibrium steady state, the 'free energy' analog  $\mathcal{F}$  in Eq. (16) can be used to define a new 'entropy' scalar as the Legendre transform

$$S \equiv \mathcal{F}(\mathbf{J}) - \mathbf{J} \cdot \nabla_{\mathbf{J}} \mathcal{F}$$
(17)

Again, a first-order dynamic equation follows using a stochastic version of the Onsager formalism from nonequilibrium thermodynamics (Groot and Mazur 1984)

$$\mathrm{d}J_t^i \approx \left(\sum_k \mu_{i,k} \partial \mathcal{S} / \partial J_t^k\right) \mathrm{d}t + \sigma_i J_t^i \mathrm{d}B_t \tag{18}$$

where  $\mu_{i,k}$  defines a diffusion matrix, the  $\sigma_i$  are parameters, and  $dB_t$  represents a noise that may be colored, i.e., not the usual Brownian motion under undifferentiated white noise.

If it is possible to factor out  $J^i$ , then Eq. (18) can be represented in the form

$$\mathrm{d}J_t^i = J_t^i \mathrm{d}Y_t^i \tag{19}$$

where  $Y_t^i$  is a stochastic process.

The expectation of J can then be found in terms of the Doleans-Dade exponential (Protter 1990) as

$$E(J_t^i) \propto \exp(Y_t^i - 1/2[Y_t^i, Y_t^i])$$
(20)

where  $[Y_t^i, Y_t^i]$  is the quadratic variation in the stochastic process  $Y_t^i$  (Protter 1990). Heuristically, invoking the mean value theorem, if

$$1/2d[Y_t^i, Y_t^i]/dt > dY_t^i/dt,$$
(21)

then the pathological ground state is stable: deviations from nonequilibrium steady state measured by  $J_t^i$  then converge in expectation to 0. That is, sufficient ongoing 'noise'—determining the quadratic variation terms—can lock-in the failure of neurodevelopment with high probability, in this model.

Parallel stability arguments arise in ecosystem resilience theory (Holling 1973) which characterizes multiple quasi-stable nonequilibrium steady states among interacting populations. Pristine alpine lake ecosystems, having limited nutrient inflows, can be permanently shifted into a toxic eutrophic state by excess nutrient influx, e.g., a sewage leak and fertilizer runoff. Once shifted, the lake ecology will remain trapped in a mode of recurrent 'red tide'-like plankton blooms even after sewage or fertilizer inflow is stemmed.

The quadratic variation in a stochastic process  $X_t$ , which we write as  $[X_t, X_t]$ , is important to understanding the pathological stability of 'eutrophic' neurodevelopmental trajectories, in this model. It can be estimated from appropriate time series

data using a Fourier expansion methodology adapted from financial engineering, as described in Sect. 6.

#### **4** The Mitochondrial Connection

Development is not simply a matter of response to external signals, powerful as such effects may be. Metabolic free energy—the high-energy conversion of ATP to ADP—powers the many cognitive process of gene expression that must control developmental trajectories. Most directly, we can posit a rate–distortion argument in which a developmental message is sent along biochemical channels, and the success of failure measured by complicated control and feedback mechanisms as indicated by the schematic of Fig. 1. In essence, there must be a parallel argument to that leading to equation (10), where T is replaced by the rate of metabolic free energy M.

Assuming a Gaussian channel, having the rate-distortion function  $R(D) = (1/2) \log[\sigma^2/D]$ , where D is the average distortion by the square measure (Cover and Thomas 2006) and  $\sigma^2$  the inherent channel noise, we can write, for the mean of  $D = \sigma^2 \exp[-2R]$ ,

$$\langle D \rangle = \frac{\int_0^\infty \sigma^2 \exp[-2R] \exp[-R/\omega M] dR}{\int_0^\infty \exp[-R/\omega M] dR}$$
$$= \frac{\sigma^2}{2\omega M + 1}$$
(22)

 $\omega$  represents the efficiency with which the system converts mitochondrial free energy into control information. Small  $\omega$  implies greatly increased levels of mitochondrial free energy are necessary for successful development, i.e., small  $\langle D \rangle$ . The obvious inference is that  $\omega$  will be affected by the degree of integrated environmental insult indexed by  $\mathcal{T}$ , so that we can to first order at least assume  $\omega = \omega_0 \mathcal{T}$  and write the synergistic relation

$$\langle D \rangle = \frac{\sigma^2}{2\omega_0 \mathcal{T} M + 1} \tag{23}$$

inversely characterizing the success of the developmental control systems: large  $\langle D \rangle$  indicates failure. Other channels, as a consequence of the convexity of the rate–distortion function in *D*, will have similar expressions.

A next level of approximation takes M itself as a monotonic increasing function of  $\mathcal{T}$ —normalized by  $\alpha_0$  of Eq. (2)—so that, to first order,  $\langle D \rangle \propto 1/\mathcal{T}^2$ . Under such a model, rising environmental insult, leading to the condition  $\mathcal{T} < 1$ , rapidly distorts developmental process by impairing mechanisms for both the generation and use of mitochondrial free energy.

#### **5** Discussion and Conclusions

Chemical exposures, pre- and neonatal infections, psychosocial stress, genetic predisposition, and the cross-generational cultural and epigenetic impacts of these and other toxicants become an integrated, perhaps synergistic, signal that can overwhelm essential neurodevelopmental regulation, demanding levels of mitochondrial free energy that cannot be met. Such insult may, as well, directly interfere with the production of mitochondrial free energy. We have characterized that dynamic through statistical models based on the asymptotic limit theorems of control and information theories, models that are the functional equivalent of the usual least-squares regression based on other asymptotic limit theorems of probability theory. The greatest scientific utility of such models remains the experimental or observational comparison of similar systems under different, and different systems under similar, conditions.

The use of such tools, however, often is not easy, as the sometimes deceptive subtleties of 'ordinary' regression remind us. Nonetheless, the conceptual approach taken here may still illuminate empirical studies.

A recent paper by Berman et al. (2016) describes the phenomenon of childhoodonset schizophrenia, a 'pure' form of the disorder observed without the often confusing correlates of an extended disease course (Lancaster and Hall 2016). Berman et al. write:

... [W]e examined large-scale network interactions in childhood-onset schizophrenia ... Using ... resting-state functional magnetic resonance imaging ... [that] identified 26 regions with decreased functional correlations in schizophrenia compared to controls ...

Lancaster and Hall (2016) find that the results of Berman et al. are compatible with a pathodevelopmental model in which patients with childhood-onset schizophrenia experience excessive 'over-pruning' of short-distance functional connections.

By contrast, autism spectrum disorders are marked by excessive early neural growth. Rapoport et al. (2009) assert that in autism there is an acceleration or excess of early postnatal brain development (1–3 years), whereas in childhood-onset schizophrenia (COS), there is exaggeration of the brain maturation processes of childhood and early adolescence (10–16 years):

Both could be seen as 'increased gain' of general developmental processes, albeit at different stages; both patterns could also be seen as an abnormal 'shift to the left' with respect to age compared to normal brain development, with autism showing initial overgrowth and COS showing greater 'pruning down' of the cortex in early and middle parts of the trajectory; both accelerations normalizing with age ...

Indeed, a recent comprehensive analysis of US insurance data indicates a strong role for environmental factors in the etiology of autism spectrum disorders. Rzhetsky et al. (2014) write

By analyzing the spatial incidence patterns of autism and intellectual disability drawn from insurance claims for nearly one third of the total US population, we found strong statistical evidence that environmental factors drive apparent spatial heterogeneity of both phenotypes [intellectual disability and autism] while economic incentives and population structure appear to have relatively large albeit weaker effects. The strongest predictors for autism were associated with the environment ... The environmental factors implicated so far include pesticides ... environmental lead ..., sex hormone analogs ... medications ... plasticizers ... and other synthetic molecules ...

It is very likely that the list of environmental factors potentially affecting development of human embryo is large and yet predominantly undocumented . . . Our results have implications for the ongoing scientific quest for the etiology of neurodevelopmental disorders. We provide evidence [for] routinely expanding the scope of inquiry to include environmental, demographic and socioeconomic factors, and governmental policies at a broad scale in a unified geospatial framework.

Environmental effects are now frequently cited as important in the etiology of autism and similar conditions (e.g., Croen et al. 2011; Landrigan 2010; DeSoto 2009). Keil and Lein (2016) in particular identify epigenetic mechanisms linking environmental chemical exposures to risk of autism spectrum disorders. Govorko et al. (2012) explore the male germline transmission of adverse effects of alcohol on fetal development.

Thus, it seems reasonable to infer that cross-generational transmission of gene methylation may also affect probabilities of ASD and other neurodevelopmental disorders. As Bohacek et al. (2013) comment,

Psychiatric diseases are multifaceted disorders with complex etiology, recognized to have strong heritable components. Despite intense research efforts, genetic loci that substantially account for disease heritability have not yet been identified. Over the last several years, epigenetic processes have emerged as important factors for many brain diseases, and the discovery of epigenetic processes in germ cells has raised the possibility that they may contribute to disease heritability and disease risk.

They specifically note '... [E]vidence suggests that highly stressful experiences at different stages of life can markedly affect behaviors across generations and might constitute heritable risk factors for affective disorders' and go on to examine the opposite effects of chemical exposures and environmental enrichment.

One central feature of the cognitive 'phase change' approach above is the possibility of a 'supercooled' state during critical neurodevelopmental periods. That is, although the 'temperature' defined by  $\mathcal{T}$  falls below threshold for phase transition, 'condensation' into nonfunctional neuronetwork configuration during a critical growth domain is made more probable rather than inevitable.

Under such a condition, however, as with supercooled liquids, some sudden perturbation can then trigger 'crystallization' from high to low symmetry states, i.e., from a normal system capable of the full 'global workspace' dynamics that Bernard Baars asserts are necessary and sufficient for consciousness in higher animals (Wallace 2012), to a fractured and fragmented structure in which essential subcomponent networks are not sufficiently linked, or become, in fact, overlinked. Different condensation dynamics would broadly account for the observations of Berman et al. and Rapoport et al., the difference between autism spectrum and COS disorders being seen as different condensation phases. Typically, in such 'spontaneous symmetry breaking,' there will be only a small number of possible different phases. Comorbidity would be seen as the existence of both possible phase types in the same individual. albeit with distinct cultural convolutions.

This inference may constitute the most central outcome of the modeling exercise, i.e., that 'environmental' stress, in a large sense, during a critical growth regime can trigger a relatively small number of characteristic phase change analogs in neurode-velopment, although the symmetry shifts will likely involve subtle groupoid changes rather than alterations of the finite groups more familiar from network theory (Yeung 2008, Ch. 16; Golubitsky and Stewart 2006). Again, simultaneous occurrence of several such 'phase condensations' would account for observed patterns of comorbidity,

Indeed, as Wallace (2015c) puts it, the stabilization of human cognition via feedback from embedding social and cultural contexts is a dynamic process deeply intertwined with it, constituting the 'riverbanks' directing flow of a stream of generalized consciousness at various scales and levels of organization: Cultural norms and social interaction are synergistic with individual and group cognition and their disorders. That analysis finds high rates of psychopathic and antisocial personality disorder, as well as obsessive/compulsive disorder, to be culture-bound syndromes particular to Western 'atomistic' societies, or to those undergoing social disintegration. Some such cultural patterning may well express itself across the forms of developmental neural malcondensation described here (e.g., Kleinman 1991; Kleinman and Cohen 1997).

While detailed application of the modeling strategies outlined here to experimental or clinical data remains to be done, the unification, after a concerted 50-year effort, of control and information theories via the data rate theorem may provide opportunity for conceptual advance. Although high-end neural structures and the genetic regulators that build them are most definitely not computers in the severely limited mathematical venue of the Turing Machine, all such systems—including computers—are bounded by the asymptotic limit theorems that constrain the generation and transmission of information in the context of dynamic control.

### **6** Mathematical Appendix

#### 6.1 An RDT Proof of the DRT

The rate-distortion theorem of information theory asks how much a signal can be compressed and have average distortion, according to an appropriate measure, less than some predetermined limit D > 0. The result is an expression for the minimum necessary channel capacity, R, as a function of D. See Cover and Thomas (2006) for details. Different channels have different expressions. For the Gaussian channel under the squared distortion measure,

$$R(D) = \frac{1}{2} \log \left[ \frac{\sigma^2}{D} \right] D < \sigma^2$$
  

$$R(D) = 0 \ D \ge \sigma^2$$
(24)

where  $\sigma^2$  is the variance of channel noise having zero mean.

Our concern is how a control signal  $u_t$  is expressed in the system response  $x_{t+1}$ . We suppose it possible to deterministically retranslate an observed sequence of system outputs  $x_1, x_2, x_3, ...$  into a sequence of possible control signals  $\hat{u}_0, \hat{u}_1, ...$  and to compare that sequence with the original control sequence  $u_0, u_1, ...$ , with the difference between them having a particular value under the chosen distortion measure, and hence an observed average distortion.

The correspondence expansion is as follows.

Feynman (2000), following ideas of Bennett, identifies information as a form of free energy. Thus, R(D), the minimum channel capacity necessary for average distortion D, is also a free energy measure, and we may define an entropy S as

$$S \equiv R(D) - DdR/dD \tag{25}$$

For a Gaussian channel under the squared distortion measure,

$$S = 1/2\log[\sigma^2/D] + 1/2$$
(26)

Other channels will have different expressions.

The simplest dynamics of such a system are given by a nonequilibrium Onsager equation in the gradient of S, (Groot and Mazur 1984) so that

$$\mathrm{d}D/\mathrm{d}t = -\mu\mathrm{d}S/\mathrm{d}D = \frac{\mu}{2D} \tag{27}$$

By inspection,

$$D(t) = \sqrt{\mu t} \tag{28}$$

which is the classic outcome of the diffusion equation. For the 'natural' channel having  $R(D) \propto 1/D$ ,  $D(t) \propto$  the cube root of *t*.

This correspondence reduction allows an expansion to more complicated systems, in particular, to the control system of Fig. 1.

Let  $\mathcal{H}$  be the rate at which control information is fed into an inherently unstable control system, in the presence of a further source of control system noise  $\beta$ , in addition to the channel noise defined by  $\sigma^2$ . The simplest generalization of Eq. (27), for a Gaussian channel, is the stochastic differential equation

$$\mathrm{d}D_t = \left[\frac{\mu}{2D_t} - M(\mathcal{H})\right]\mathrm{d}t + \beta D_t \mathrm{d}W_t \tag{29}$$

where  $dW_t$  represents white noise and  $M(\mathcal{H}) \ge 0$  is a monotonically increasing function.

This equation has the nonequilibrium steady-state expectation

$$D_{\rm nss} = \frac{\mu}{2M(\mathcal{H})} \tag{30}$$

measuring the average distortion between what the control system wants and what it gets. In a sense, this is a kind of converse to the famous radar equation which states that a returned signal will be proportional to the inverse fourth power of the distance between the transmitter and the target. But there is a deeper result, leading to the DRT. Applying the Ito chain rule to Eq. (29) (Protter 1990; Khashminskii 2012), it is possible to calculate the expected variance in the distortion as  $E(D_t^2) - (E(D_t))^2$ . But application of the Ito rule to  $D_t^2$  shows that no real number solution for its expectation is possible unless the discriminant of the resulting quadratic equation is  $\ge 0$ , so that a necessary condition for stability is

$$M(\mathcal{H}) \ge \beta \sqrt{\mu}$$
  
$$\mathcal{H} \ge M^{-1}(\beta \sqrt{\mu})$$
(31)

where the second expression follows from the monotonicity of M.

As a consequence of the correspondence reduction leading to Eq. (29), we have generalized the DRT of Eq. (2). Different 'control channels,' with different forms of R(D), will give different detailed expressions for the rate of generation of 'topological information' by an inherently unstable system.

#### 6.2 A Black–Scholes Model

We look at  $\mathcal{H}(\rho)$  as the control information rate 'cost' of stability at the integrated environmental insult  $\rho$ . To determine the mathematical form of  $\mathcal{H}(\rho)$  under conditions of volatility i.e., variability proportional to a signal, we must first model the variability of  $\rho$ , most simply taken as

$$d\rho_t = g(t, \rho_t)dt + b\rho_t dW_t \tag{32}$$

Here,  $dW_t$  is white noise and—counterintuitively—the function  $g(t, \rho)$  will fall out of the calculation on the assumption of certain regularities.

 $\mathcal{H}(\rho_t, t)$  is the minimum needed incoming rate of control information under the data rate theorem. Expand  $\mathcal{H}$  in  $\rho$  using the Ito chain rule (Protter 1990):

$$d\mathcal{H}_{t} = [\partial \mathcal{H}/\partial t + g(\rho_{t}, t)\partial \mathcal{H}/\partial \rho + \frac{1}{2}b^{2}\rho_{t}^{2}\partial^{2}\mathcal{H}/\partial \rho^{2}]dt + [b\rho_{t}\partial \mathcal{H}/\partial \rho]dW_{t}$$
(33)

It is now possible to define a Legendre transform, L, of the rate  $\mathcal{H}$ , by convention having the form

$$L = -\mathcal{H} + \rho \partial \mathcal{H} / \partial \rho \tag{34}$$

 $\mathcal{H}$  is an information index, a free energy measure in the sense of Feynman (2000), so that L is a classic entropy measure.

We make an approximation, replacing dX with  $\Delta X$  and applying Eq. (33), so that

$$\Delta L = \left(-\partial \mathcal{H}/\partial t - \frac{1}{2}b^2\rho^2\partial^2 \mathcal{H}/\partial\rho^2\right)\Delta t \tag{35}$$

According to the classical Black–Scholes model (Black and Scholes 1973), the terms in g and  $dW_t$  'cancel out,' and white noise has been subsumed into the Ito

correction factor, a regularity assumption making this an exactly solvable but highly approximate model.

The conventional Black–Scholes calculation takes  $\Delta L/\Delta T \propto L$ . At nonequilibrium steady state, by some contrast, we can assume  $\Delta L/\Delta t = \partial \mathcal{H}/\partial t = 0$ , giving

$$-\frac{1}{2}b^2\rho^2\partial^2\mathcal{H}/\partial\rho^2 = 0$$
(36)

so that

$$\mathcal{H} = \kappa_1 \rho + \kappa_2 \tag{37}$$

The  $\kappa_i$  will be nonnegative constants.

#### 6.3 Estimating the Quadratic Variation from Data

So-called white noise has quadratic variation  $\propto t$ . The 'colored' noise relation can be estimated from the observed periodogram using the methods of Dzhaparidze and Spreij (1994).

For a stochastic process  $X_t$  and a finite stopping time T and each real number  $\lambda$ , the *periodogram* of X evaluated at T is defined as

$$I_T(X;\lambda) \equiv \left|\int_0^T \exp[i\lambda t] \mathrm{d}X_t\right|^2 \tag{38}$$

Take  $\epsilon$  as a real random variable that has a density  $\omega$  symmetric around zero and consider, for any positive real number L, the quantity

$$E_{\epsilon}[I_T(X; L\epsilon)] = \int_{-\infty}^{+\infty} I_T(X; Ls)\omega(s) ds$$
(39)

Dzhaparidze and Spreij (1994) show that, for  $L \to \infty$ ,

$$E_{\epsilon}[I_T(X; L\epsilon)] \to [X_T, X_T] \tag{40}$$

Thus, the quadratic variation can be statistically estimated from observational time series data, as is routinely done in financial engineering, from which, in fact, this analysis is taken.

Acknowledgments The author thanks Drs. D.N. Wallace, J.L. Rapoport, K. Keil, and P. Lein for useful discussions and several anonymous reviewers for comments helpful in revision.

#### References

Atlan H, Cohen I (1998) Immune information, self-organization, and meaning. Int Immunol 10:711–717 Ben-Shachar D (2002) Mitochondrial dysfunction in schizophrenia: a possible linkage to dopamine. J Neurochem 83:1241–1251

- Berman R, Gotts S, McAdams HJ, Greenstein D, Lalonde F, Classen L, Watsky R, Shora L, Ordonex A, Raznahan A, Martin A, Gogtay N, Rapoport J (2016) Disrupted sensorimotor and social–cognitive networks underlie symptoms in childhood-onset schizophrenia. Brain 139:276–291
- Black F, Scholes M (1973) The pricing of options and corporate liabilities. J Polit Econ 81:637–654
- Bohacek J, Gapp K, Saab B, Mansuy I (2013) Transgenerational epigenetic effects on brain functions. Biol Psychiatry 73:313–320
- Clay H, sillivan S, Konradi C (2011) Mitochondrial dysfunction and pathology in bipolar disorder and schizophrenia. Int J Dev Neurosci 29:311–324
- Corbin J, Gaiano N, Juliano S et al (2008) J Neurochem 106:2272-2287
- Cover T, Thomas J (2006) Elements of information theory, 2nd edn. Wiley, New York
- Croen L, Grether J, Yoshida C, Odouli R, Hendrick V (2011) Antidepressant use during pregnancy and childhood autism spectrum disorders. Arch Gen Psychiatry 68:1104–1112
- de Groot S, Mazur P (1984) Nonequilibrium thermodynamics. Dover, New York
- DeSoto M (2009) Ockham's razor and autism: the case for developmental neurotoxins contributing to a disease of neurodevelopment. Neurotoxicology 30:331–337
- Dzhaparidze K, Spreij P (1994) Spectral characterization of the optional quadratic variation. Stoch Process Appl 54:165–174
- Feynman R (2000) Lectures on computation. Westview Press, Boulder
- Garcez P, Loiola E Correia, da Costa RM et al (2016) Zika virus impairs growth in human neurospheres and brain organoids. https://peerj.com/preprints/1817v3
- Giulivi C et al (2010) Mitochondrial dysfunction in autism. J Am Med Assoc 304:2389–2395
- Glazebrook JF, Wallace R (2009) Rate distortion manifolds as model spaces for cognitive information. Informatica 33:309–346
- Goh SA, Dong Z, Zhang Y, DiMauro S, Peterson B (2014) Mitochondrial dysfunctions as a neurobiological subtype of autism spectrum disorder. JAMA Psychiatry 71:665–671. doi:10.1001/jamapsychiatry. 2014.179
- Golubitsky M, Stewart I (2006) Nonlinear dynamics of networks: the groupoid formalism. Bull Am Math Soc 43:305–364
- Govorko D, Bekdash R, Zhang C, Sakar D (2012) Male germline transmits fetal alcohol adverse effect on hypothalamic proopieomelanocortin gene across generations. Biol Psychiatry 72:378–388
- Hayden E (2016) Zika highlights role of controversial fetal-tissue research: fetal tissue may prove crucial to probing link between virus and brith defects. Nature 532:16
- Holling C (1973) Resilience and stability of ecological systems. Annu Rev Ecol Syst 4:1-23
- Keil K, Lein P (2016) DNA methylation: a mechanism linking environmental chemical exposures to risk of autism spectrum disorders? Environ Epigenet 2:dvv012. doi:10.1093/eep/dvv012
- Khashminskii R (2012) Stochastic stability of differential equations, 2nd edn. Springer, New York
- Kleinman A (1991) Rethinking psychiatry: from cultural category to personal experience. Free Press, New York
- Kleinman A, Cohen A (1997) Psychiatry's global challenge. Sci Am 276(3):86-89
- Lancaster T, Hall J (2016) Altered intra- inter-network dynamics reflect symptom dimensions in childhoodonset schizophrenia. Brain 139(10):12
- Landrigan P (2010) What causes autism? Exploring the environmental contribution. Curr Opin Pediatr 22:219–225
- Nair G, Fagnani F, Zampieri S, Evans R (2007) Feedback control under data rate constraints: an overview. Proc IEEE 95:108–137
- Palmieri L, Persico A (2010) Mitochondrial dysfunction in autism spectrum disorders: cause or effect? Biochim Biophys Acta 1797:1130–1137
- Pettini M (2007) Geometry and topology in Hamiltonian dynamics. Springer, New York
- Prabakaran S et al (2004) Mitochondrial dysfunction in schizophrenia: evidence for compromised brain metabolism and oxidative stress. Mol Psychiatry 9:6840697
- Protter P (1990) Stochastic integration and differential equations. Springer, New York
- Rapoport J, Chavez A, Greenstein D, Addington A, Gogtay N (2009) Autism-spectrum disorders and childhood onset schizophrenia: clinical and biological contributions to a relationship revisited. J Am Acad Child Adolesc Psychiatry 48:10–18
- Rapoport J, Giedd J, Gogtay N (2012) Neurodevelopmental model of schizophrenia: update 2012. Mol Psychiatry 17:1228–1238

- Rossignol R, Frye R (2010) Mitochondrial dysfunction in autism spectrum disorders: a systematic review and meta-analysis. Mol Psychiatry 17:290–314
- Rzhetsky A, Bagley S, Wang K, Lyttle C, Cook E, Altman R, Gibbons R (2014) Environmental and statelevel regulatory factors affect the incidence of autism and intellectual disability. PLOS Comput Biol 10(3):e1003581
- Scaglia F (2010) The role of mitochondrial dysfunction in psychiatric disease. Dev Disabil Res Rev 16:136– 143
- Shao L et al (2008) Mitochondrial involvement in psychiatric disorders. Trends Mol Med 40:281-295
- Tiberi L, Vanderhaeghen P, van den Ameele J (2012) Cortical neurogenesis and morphogens: diversity of cues, sources and functions. Curr Opin Cell Biol 24:269–276
- van Os J, Kenis G, Rutten B (2010) The environment and schizophrenia. Nature 468:203-212
- Wallace R, Wallace D (2010) Gene expression and its discontents: the social production of chronic disease. Springer, New York
- Wallace R (2012) Consciousness, crosstalk, and the mereological fallacy: an evolutionary perspective. Phys Life Rev 9:426–453
- Wallace R (2015a) An information approach to mitochondrial dysfunction: extending Swerdlow's hypothesis. World Scientific, Singapore
- Wallace R (2015b) Metabolic free energy and biological codes: a 'Data Rate Theorem' aging model. Bull Math Biol 77:879–903
- Wallace R (2015c) Closed-system 'economic' models for psychiatric disorders: Western atomism and its culture-bound syndromes. Cogn Process 16:279–290
- Wallace R (2016) Subtle noise structures as control signals in high-order biocognition. Phys Lett A 380:726– 729
- Weinstein A (1996) Groupoids: unifying internal and external symmetry. Not Am Math Assoc 43:744–752 Yeung R (2008) Information theory and network coding. Springer, New York