

Advancing System Performance with Redundancy: From Biological to Artificial Designs

Anh Tuan Nguyen

nguy2833@umn.edu

Jian Xu

xuxx1268@umn.edu

Diu Khue Luu

luu00009@umn.edu

Biomedical Engineering, University of Minnesota, Minneapolis, MN 55455, U.S.A.

Qi Zhao

qzhao@umn.edu

Computer Science and Engineering, University of Minnesota, Minneapolis,

MN 55455, U.S.A.

Zhi Yang

yang5029@umn.edu

Biomedical Engineering, University of Minnesota, Minneapolis, MN 55455, U.S.A.

Redundancy is a fundamental characteristic of many biological processes such as those in the genetic, visual, muscular, and nervous systems, yet its driven mechanism has not been fully comprehended. Until recently, the only understanding of redundancy is as a mean to attain fault tolerance, which is reflected in the design of many man-made systems. On the contrary, our previous work on redundant sensing (RS) has demonstrated an example where redundancy can be engineered solely for enhancing accuracy and precision. The design was inspired by the binocular structure of human vision, which we believe may share a similar operation. In this letter, we present a unified theory describing how such utilization of redundancy is feasible through two complementary mechanisms: representational redundancy (RPR) and entangled redundancy (ETR). We also point out two additional examples where our new understanding of redundancy can be applied to justify a system's superior performance. One is the human musculoskeletal system (HMS), a biological instance, and the other is the deep residual neural network (ResNet), an artificial counterpart. We envision that our theory would provide a framework for the future development of bio-inspired redundant artificial systems, as well as assist studies of the fundamental mechanisms governing various biological processes.

1 Introduction

Redundancy is a well-known characteristic of many biological processes from the molecular to the systematic level. For example, the human's genome is highly redundant: a particular gene can be duplicated at various regions of DNA, and multiple genes can encode the same or similar biochemical functions and phenotype expressions. This genetic redundancy and functional redundancy is observed in many crucial pathways of the developmental, signaling, and cell cycle processes (Tautz, 1992; Nowak, Boerlijst, Cooke, & Smith, 1997; Kafri, Springer, & Pilpel, 2009). High levels of redundancy are also present in the nervous system. While observing the behavioral resistance and recovery from massive damage, Glassman (1987) speculated that the human brain has evolved with "spare capacity"—at least twice the size as necessary for its basic function. Subsequent studies using both computational and empirical approaches have shown that redundancy is evident not only in the neuronal circuits' architecture and interconnections, as suggested by Glassman (i.e., physical redundancy), but also in the way neuron populations encode, retrieve, and manipulate information (i.e., information redundancy) (Panzeri, Schultz, Treves, & Rolls, 1999; Narayanan, Kimchi, & Laubach, 2005; Averbeck, Latham, & Pouget, 2006; Pitkow & Angelaki, 2017). There is no doubt that neural redundancy is one of the driving forces allowing the brain to facilitate complex processes of learning, memorizing, and self-repairing.

In many scenarios, the redundant structure of a biological system can be seen as a consequence of the evolutionary process. Under the pressure of natural selection, living organisms develop multiple strategies that achieve the same goal: survival. It is not uncommon for distinct strategies that emerge from entirely different evolutionary pathways to resolve the same biological problem. These strategies could coexist in the same ecosystem or even the same organism's genome, creating observable repeated evolutionary behaviors such as functional redundancy, parallel evolution, and convergent evolution (York & Fernald, 2017). Redundancy could also be driven by the defense against failures, which contributes to an overall higher survival rate. For example, gene duplication has been shown to mitigate effects of mutations and reduce the chance of catastrophic phenotype expression (Kafri et al., 2009). Redundancy also helps the human brain tolerate significant damage and loss of mass due to injury or disease. Damaged neurons and brain tissue generally do not regrow, yet their redundant structures allow reorganization of the neuronal circuits to recover many basic brain functions (Glassman, 1987). Redundancy also increases the organism's adaptability. For example, genetic and functional duplication has been shown to be the basis of phenotypic plasticity, which allows an organism to adapt and survive rapidly changing endogenous and exogenous environmental conditions (Kafri et al., 2009).

Many of these principles find application in the design and engineering of artificial systems. However, almost all intentional utilization of redundancy in man-made systems focuses on enhancing reliability, the importance of which is often overshadowed by the system's performance. Also, methods for incorporating redundancy involve the replication of partial or entire systems, which require large amounts of overhead. As a result, redundant designs such as dual modular redundancy (DMR) and triple modular redundancy (TMR) are mostly found in specialized systems that perform critical functions, such as aircraft controllers, biomedical implants, and computer servers.

In this letter, we argue two counterintuitive arguments. First, redundancy can be engineered solely for enhancing systems' performance regarding accuracy and precision instead of reliability and plasticity. Second, a practical implementation of information redundancy is feasible without replication and excessive resource overhead or physical redundancy, thus mitigating trade-offs encountered by conventional designs. The performance boost in our proposed framework is achieved by employing two complementary mechanisms: representational redundancy (RPR) and entangled redundancy (ETR). RPR describes how information is redundantly encoded and processed, while ETR allows realizing an RPR scheme in actual applications.

We have shown Nguyen, Xu, and Yang (2015, 2016) a simple but practical application where a redundant sensing (RS) architecture resembling the binocular structure of the human vision is applied to enhance the precision of a man-made sensor without incurring compromises often seen in conventional architectures. In theory, the RPR and ETR principles utilized in our design can be generalized to different applications and also serve as a fundamentally structural characteristic of more complex systems. We assert this argument further in this letter by examining empirical evidence in two different systems from two distinct fields of science and engineering. One is the HMS: a biological system where redundancy contributes to generating complex and precise muscle movements. The other is the ResNet: an artificial deep learning architecture where redundancy helps accomplish superior predicting accuracy compared to conventional methods. By understanding the subtle yet sophisticated roles of redundancy in these systems, we believe that the findings would not only enrich our knowledge of biological processes but also inform the derivation of new methods for advancing the performance of man-made designs.

The remainder of the letter is organized as follows. Section 2 consolidates our redundant model of RPR and ETR mechanisms. Section 3 examines the evidence suggesting the implication of our model in biological and artificial systems, which includes the proposed sensor design, the HMS, and the ResNet. Section 4 concludes and offers discussions on the future development of the proposed theory. The appendix summarizes the new terminologies and definitions used in this letter.

2 Advancing Performance with Redundancy

2.1 Representational Redundancy (RPR). The vast majority of artificial systems are designed on an orthogonal scheme of information representation where each entry of information is encoded by a unique configuration of the system. An entry of information can be an input value, a desirable output, an intermediate instance, or an operation of the information processing pathway. Such orthogonal systems excel in efficiency because they allow rapidly and unambiguously acquiring, processing, and storing of information. However, any encoding and decoding scheme in practice suffers from an inevitable level of error resulting in the limitation of its accuracy. In many computational models, this limitation is described by Shannon's (1948) theorem. Because of the uniqueness of the representation scheme, any error acquired during the sampling, processing, and storing of information cannot be easily corrected without an overhead in term of resources such as power, bandwidth, and memory.

The RPR concept is designed to overcome conventional limitations by embracing a nonorthogonal scheme of information representation. Subsequently, every entry of information can be encoded by numerous distinct system configurations, including the conventional one. These configurations are referred to as the system's microstates. If the microstates are designed such that their responses to error are nonhomologous, in any given instance, provided a sufficient number of distinct microstates, there exist with asymptotic certainty one or more microstates that have a smaller error than the conventional representation. Therefore, an overall RPR system would have a theoretical accuracy almost always superior to the conventional counterpart with similar structure provided the optimal microstates for every entry of information can be identified.

2.2 Entangled Redundancy (ETR). The number of microstates represents information capacity—an abstract property of the design that is not necessarily proportional to its physical size. In order to effectively deploy an RPR system in practice, the microstates must be designed so that they do not incur excessive resource overhead. In other words, information redundancy must be achieved without physical redundancy. As a result, the statistical distribution of the microstates with respect to error cannot be independent; it must be partially correlated or entangled. This concept is known as ETR. The level of entanglement should be engineered sufficiently to create excessive redundancy without trading off large amounts of resources. ETR should be differentiated from the conventional method of creating redundancy by replication where the distribution of repeated instances is independent of each other and the resource utilization is linearly proportional to the level of redundancy.

Figure 1 illustrates the architectural distinction among three information processing systems: a conventional orthogonal system (COS), a conventional redundant system (CRS), and a proposed representational

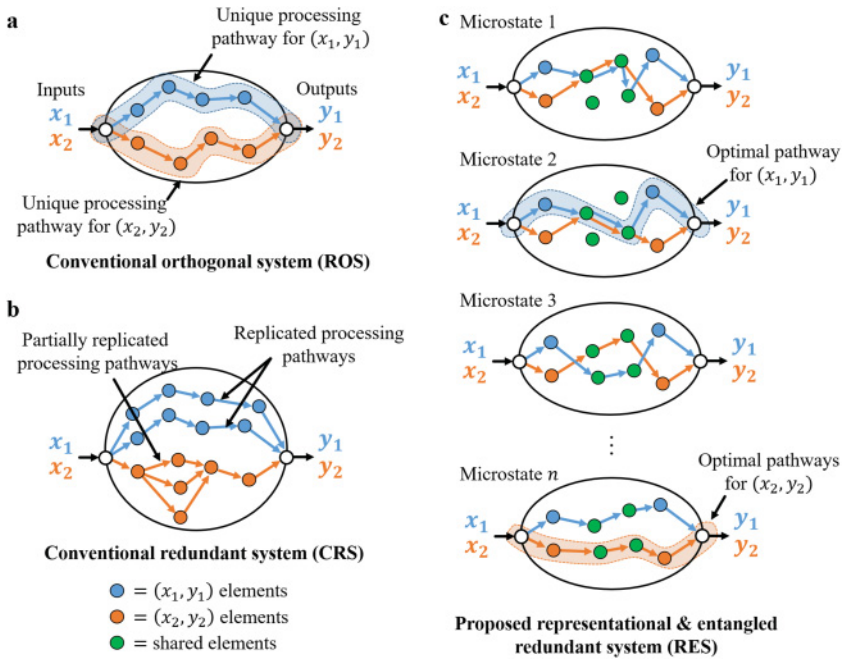


Figure 1: The architectural differences among three information processors: a conventional orthogonal system (COS), a conventional redundant system (CRS), and a proposed representational and entangled redundant system (RES). An entry of information in this example is a processing pathway that takes an input x_i and produces a corresponding output y_i ($i = 1, 2, \dots$). (a) In the COS, every (x_i, y_i) pair is represented by a unique pathway. (b) In the CRS, the pathways are partially or entirely replicated, which gives the system fault-tolerance properties and a marginal accuracy gain. (c) In the RES, because of the entanglement among different processing pathways, the system can be configured to various microstates with a distinctive amount of error while utilizing the same number of elements. An exponential level of information redundancy is effectively realized with minimal physical redundancy. Major accuracy enhancement is feasible if the microstates with the least error can be found for every (x_i, y_i) pair.

and entangled redundant system (RES). In these processors, an entry of information is a pathway that takes an input x_i and produces a corresponding output y_i ($i = 1, 2, \dots$). In the COS (see Figure 1a), every input-output pair (x_i, y_i) is represented by a unique pathway that has a determined error that cannot be easily removed without physically altering the system. The pathways in the CRS (see Figure 1b) are partially or entirely replicated, which requires a proportional resource overhead (i.e., physical redundancy). Although in practice, the replication is mostly used for fault tolerance, a marginal accuracy gain is feasible by selecting the pathway with

the least error for each input instance. The RES (see Figure 1c) incorporates redundancy by having the pathways of different (x_i, y_i) pairs share certain processing elements. While the total number of elements (shared and unshared) remains the same, each (x_i, y_i) pathway in the RES can now be represented by various different systems' configurations or microstates, which increase exponentially with the number of shared elements. In other words, the RES achieves an exponential level of information redundancy with minimal physical redundancy. Because each microstate has a different level of error, if the microstates with the least amount of error can be found for every (x_i, y_i) pair, major accuracy enhancement is feasible without the need to physically alter the system. In theory, the RES is superior compared to COS in terms of accuracy because there almost always exists a pathway with a lower error for any given (x_i, y_i) pair. The RES is also superior compared to CRS because an exponential level of redundancy can be achieved with minimal additional resources.

2.3 Challenges. A proper implementation of RPR and ETR in the same architecture is essential to achieve performance boosts. The goal is to create more microstates than seem needed while utilizing their entanglement to allow the microstates to coexist in superposition, thus requiring minimal additional resources (see Figure 1c). Unfortunately, there is no universal solution that can be applied to all types of systems. In our proof-of-concept system described in (Nguyen et al., 2015, 2016), redundancy is realized by integrating two similar binary-weighted arrays, whose structure resembles the human's binocular vision. In the subsequent sections, the HMS and ResNet provide additional examples where redundancy elegantly emerges in entirely different ways.

Furthermore, while a RES provides a redundant nonorthogonal structure of information representation, there is no universal solution to identify the optimal microstate given a particular input. In fact, in almost all examples of RES, it appears to be an NP-optimization problem that can be resolved only by the mean of approximation. Biological processes such as the visual and musculoskeletal systems overcome this challenge by harnessing the computational capacity of the nervous system, which is exceptionally good at approximation. A similar mechanism could be utilized by the ResNet, itself a neural network. For engineering systems such as our redundant sensor (Nguyen et al., 2015), an approximation method that consists of a one-shot unsupervised error estimation and a simplified calibration algorithm needs to be derived.

3 From Biological to Artificial Systems

3.1 Redundant Sensing and Binocular Vision. Nguyen et al. (2015) shows a proof-of-concept implementation of the system with both RPR and

ETR properties. The design called for a redundant analog-to-digital converter (ADC), a fundamental component of many sensory data acquisition systems. An entry of information is a digital code $x_D \in \{0, 1, \dots, 2^N - 1\}$ (N : resolution) representing an input analog voltage. In practice, each code is generated by assembling a set of components that are often miniature capacitors embedded on a silicon chip. The number of unit components is proportional to the required physical resources and cost. The random error that occurred during the fabrication process of these unit components (i.e., mismatch error) has been shown to be a major factor limiting the device's accuracy.

Figure 2a compares a conventional and the proposed redundant sensing (RS) architecture in a simplified case of $N = 3$. The conventional system uses a binary-weighted set of components, which is the most efficient encoding scheme yet vulnerable to mismatch error. The proposed RS architecture employs a nonorthogonal component set that satisfies both RPR and ETR requirements. With the same number of unit components ($2^N - 1 = 7$), RS allows each digital code to be generated by multiple different component assemblies (i.e., microstates). Figure 2b shows a similar concept applied to an $N = 10$ bits device with $2^N - 1 = 1023$ unit components. We examine two classes of component set design, c_1 and c_2 . The mathematical formulation of these designs can be found in Nguyen et al. (2016). The number of microstates that represents each digital code increases exponentially with N , which results in an excessive level of redundancy. Figure 2c presents simulation results of the overall system error using the Monte Carlo method ($n = 10^4$) with the error distribution of the unit components as a prior. At each digital code, the component assembly with the least amount of error is found by exhaustive search. The data demonstrate that the RS technique can substantially suppress the error in the system, leading to major precision enhancement. Moreover, the effectiveness of the error reduction is correlated with the level of redundancy: the more microstates that represent the same code, the less error can be attained. More detailed mathematical formulations as well as benchmarking of an actual device have been reported in Nguyen et al. (2016).

Interestingly, our design of the RS architecture was inspired by the binocular structure of the human visual system Nguyen et al. (2016). An RS component set resembles exchanging and integrating the information between two smaller conventional binary-weighted subarrays, which echoes the way we humans coordinate our two eyes. Thus, we ask whether RPR and ETR are also fundamental properties that facilitate visual acuity. The spatial distribution of photoreceptors on the retina is notably irregular, echoing the impact of mismatch error. This could result in the object to be registered differently by two eyes, causing acute distortion, as illustrated in Figures 3a and 3b. How does the brain compensate for these errors? Our hypothesis is that by integrating the information content obtained from both eyes, pixel

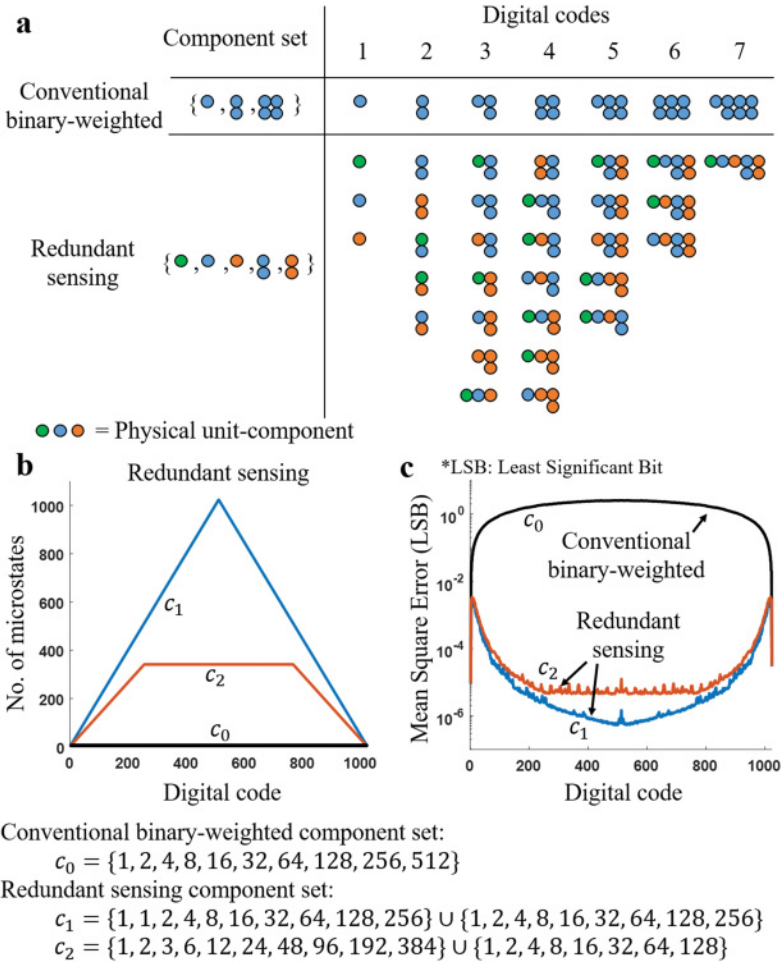


Figure 2: (a) Illustration of the RPR and ETR properties of the proposed redundant sensing (RS) architecture (Nguyen et al., 2016) for a simple system of $N = 3$ bits resolution. The device converts an analog input to a digital output by assembling a set of physical unit components. While utilizing the same number of unit components ($2^N - 1 = 7$) as a conventional binary-weighted design, the RS architecture allows each digital code to be created by multiple different assemblies (i.e., microstates). By selecting the microstate with the least error for every code, a significant boost in accuracy can be achieved. (b) A similar concept applied to an $N = 10$ bits device with two different classes of component set design. The microstate counts for each digital code increase exponentially with N . (c) Using Monte Carlo simulations ($n = 10^4$) with the error distribution of the unit components as a prior, we show that the RS technique can substantially suppress the error and enhance the overall precision of the system.

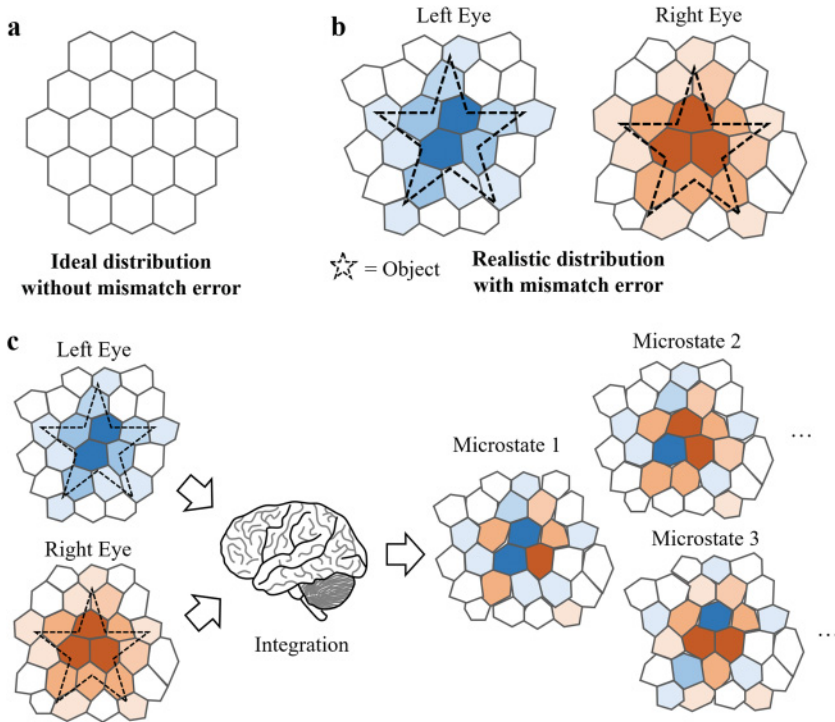


Figure 3: (a) In an ideal 2D quantizer, its pixels should uniformly distribute across the sample space without any mismatch error. (b) The spatial distribution of photoreceptors on the retina of the human eye is notably irregular. This could result in the object to be registered differently on two eyes, causing acute distortion. (c) Our hypothesis is that by integrating the information content obtained from both eyes, pixel by pixel, the brain effectively uses the binocular structure to create a massive number of representations (i.e., microstates) of the image. By means of heuristic approximation, it is possible to find the near-optimal microstate to produce the image with less distortion that we perceive.

by pixel,¹ the brain effectively utilizes the binocular structure to create a massive number of representations (i.e., microstates) of the image, as illustrated in Figure 3c. Then, through heuristic approximation means similar to our RS sensor, it is possible to find near-optimal microstates to produce the image with less distortion than we perceive. The binocular structure plays

¹Here we do not define what is considered as a pixel. It may be intuitive to think of pixels as photoreceptors, but to the brain, the smallest unit of visual information could also be the retinal ganglion cell or a subcircuit of the visual pathway.

an important role in this mechanism because it creates a form of static redundancy, allowing the brain to collect sufficient information to remedy the error. In fact, the binocular vision has been shown to help in the differentiation of fine details, even exceeding the diffraction limit of the photoreceptors, a phenomenon known as hyperacuity (Beck & Schwartz, 1979).

Furthermore, we hypothesize that the proposed redundant mechanism and computational limitation contribute to the fact that human and many higher-order animals have only two eyes. Figure 3c implies that the number of microstates, which correlates to the level of redundancy and the amount of processing power required, increases exponentially with respect to the number of eyes. Two eyes happen to be the fewest required to effectively realize such a redundant structure. Yet having two eyes has already created a substantial amount of information that visual processing directly or indirectly accounts for 30% to 60% of brain mass. Any additional eyes would overwhelm the computational capacity of the human brain. Of course, two eyes are also the minimum requirement for depth perception; however, that does not explain why no “trinocular” or “multinocular” animals exist.²

Furthermore, as a complement to the binocular structure, we conjecture that eyes’ microfixational movement or microsaccade (Martinez-Conde, Otero-Millan, & Macknik, 2013) create a form of dynamic redundancy. During microsaccades, the field of vision of each eye is sampled multiple times by different spatial configurations of photoreceptors, which resemble entangled redundant microstates and facilitate visual acuity. This observation is supported by experiments with human subjects (Hicheur, Zozor, Campagne, & Chauvin, 2013) and mathematical modeling (Hennig & Worgotter, 2004) where microsaccades have been shown to play an important role in visual precision and could lead to hyperacuity.

3.2 Muscle Redundancy. The HMS has more muscles and joints than the necessary mechanical degrees of freedom even though they are energetically expensive to produce and maintain. This paradoxical phenomenon of muscle redundancy (MR), first formulated by Bernstein (1967), presents a long-standing problem in human kinesiology of understanding how and why the human brain coordinates all muscles and joints to achieve complex movements with precision. By examining this biological process from the perspective of our model, we hope to unravel the principles underlying the behavior of MR.

A conventional interpretation would suggest that redundancy contributes to the reliability of the HMS, allowing compensation for the loss or dysfunction of individual muscles. However, emerging empirical

²Many spiders have six to eight eyes; however, it is evident that they do not have complex visual processing like higher-order animals do. The additional eyes function independently from each other and simply to increase the field of view.

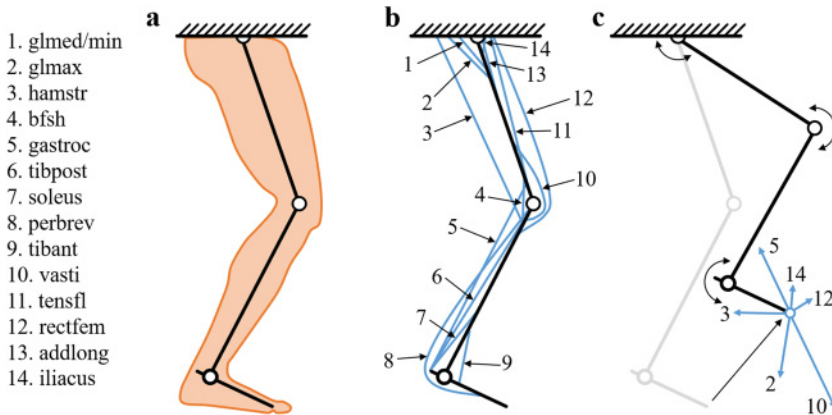


Figure 4: (a, b) A sagittal view of a human leg's mechanical model consisting of 14 muscles and muscle groups (Kutch & Valero-Cuevas, 2011). (c) Any specific movement trajectory and force can be achieved by multiple distinct muscle and joint combinations. Also, different muscles have overlapping but not exclusive mechanical functions, and all contribute with different degrees in generating force and movement. These characteristics resemble an RES with both RPR and ETR properties.

evidence implies this is not true. Even a mild dysfunction of a few critical muscles due to disorders, injuries, or aging can significantly weaken the force production and overall functions of the whole HMS (Forssberg et al., 1991; Schreuders, Selles, Roebroek, & Stam, 2006). The results are supported by Kutch and Valero-Cuevas (2011) and Valero-Cuevas (2015). Using both computational models and empirical experiments with cadaver specimens, the authors point out that less than 5% of the feasible forces and movements in their models are robust to a loss of any muscle, so it is clear that reliability is not an inherited characteristic of MR.

Instead, the redundant characteristics of the HMS resemble that of the RPR and ETR properties. Figure 4 presents an example of a component belonging to the HMS: the sagittal view of a human leg's mechanical model used by Kutch and Valero-Cuevas (2011), which consists of 14 muscles and muscle groups.³ At the kinematic and muscular levels (see Figure 4c), any

³The 14 muscles/muscle groups and their abbreviation: (1) gluteus medialis and minimus (glmed/min); (2) gluteus maximus (glmax); (3) semimembranoseus, semitendinosus and biceps femoris long head (hamstr); (4) biceps femoris short head (bfsh); (5) medial and lateral gastrocnemius (gastroc); (6) tibialis posterior (tibpost); (7) soleus (soleus); (8) peroneus brevis (perbrev); (9) tibialis anterior (tibant); (10) vastus intermedius, lateralis and medialis (vasti); (11) tensor fascia lata (tensfl); (12) rectus femoris (rectfem); (13) adductor longus (addlong); and (14) iliacus (iliacus).

specific movement trajectory and force can be achieved by numerous combinations of muscles and joints. Similarly, at the control level, each muscle consists of numerous units that can be activated by different motor neurons and patterns while resulting in the same behavior. These attributes clearly echo an RPR system where the entry of information is a specific leg movement and the microstates are different activation profiles of the muscles and joints. Furthermore, the muscles have overlapping but not exclusive mechanical functions, and all muscles contribute with different degrees in generating force and movement (Valero-Cuevas, Cohn, Yngvason, & Lawrence, 2015). These attributes are consistent with an ETR system where the mechanical entanglement allows virtually infinite configurations and dynamics to be realized with a reasonable number of muscles and joints.

These observations resonate with a number of studies where redundant characteristics have been shown to play an important role in enhancing the accuracy and precision of movements. Cleather and Bull (2010) examine two different muscular models, Delp and Horsman, in predicting the patellofemoral force during standing, jumping, and weightlifting. They conclude that the higher level of redundancy in the Horsman model contributes to its higher predictive accuracy and closer realistic approximation in all activities. The authors' conjecture is consistent with our theory, which implies that redundancy effectively increases the variability and number of independent musculoskeletal movements (i.e., microstates), so an optimal solution is more likely to be found. The argument is strengthened in Moissenet, Cheze, and Dumas (2016), where an increased level of redundancy correlates to better predicting of the accuracy of tibiofemoral contact forces in all gait patterns.

However, Valero-Cuevas et al. have a different interpretation of HMS's behaviors that perhaps, are not redundant after all. This is notably illustrated in Hagen and Valero-Cuevas (2017), where the authors examine the sagittal-plane model of the arm and find that even similar trajectories have large differences in the eccentric and concentric muscle velocities, and in Marjaninejad and Valero-Cuevas (2018), where the authors establish a formal mathematical approach to the control of tendons for anthropomorphic robots and suggest that vertebrates merely have sufficient muscles to meet the physical constraints for ecological functions. Does Valero-Cuevas's conclusion contradict our theory on redundancy? Upon closer examination, the two models are actually describing the same phenomenon.

One way to resolve this dilemma is to realize that a system can be architecturally redundant by design, while its end behavior in the real world is not. In our RS paradigm and Bernstein's classic model, the system appears to be redundant, as one kinematic outcome is encoded by many muscle and joint configurations because none of the realistic, nonideal, or random factors are considered. This "redundant" state has been shown to be a highly unstable equilibrium in our analysis. The redundancy breakdowns in the presence of even a minuscule random element (e.g. mismatch error) result

in the discrepancy in the end behavior among all the “mechanically equivalent” system configurations. These nonredundant end behaviors are what are observed in Valero-Cuevas’s model, where different kinematic trajectories express a diverse range of characteristics, including distinct eccentric and concentric muscle velocities.

Hence, the seeming dilemma between the RS/Bernstein’s model and Valero-Cuevas’s is essentially the difference in point of view. While our model looks at a system from the top-down perspective (architecture), Valero-Cuevas’s model analyzes it from the bottom-up perspective (end behavior). Both describe the same phenomenon: when realistic factors are considered, (1) different “redundant” kinematic trajectories of the HMS elicit fine discrepancy in their behavior; (2) these fine differences are recognized by the brain by their distinct eccentric and concentric muscle velocities, which ultimately result in different proprioceptive feedback signals; and (3) the ability to consistently identify and execute these fine-detailed actions greatly contributes to the capability of the the muscular system to adjust its output, generating more complex and accurate motions. In fact, Valero-Cuevas’s mathematical model indicates that adding more degrees of freedom (i.e., muscles or joints) increases functional versatility as it extends the system’s architectural redundancy. This also explains why very precise movements such as a perfect free kick of elite athletes take years to practice and refine because it is essentially a NP-hard optimization problem and increasing the redundancy results only in a further increase of complexity (Cleather & Bull, 2010; Moissenet et al., 2016).

A remaining open question is how the brain selects a specific motor control pattern among virtually infinite possibilities. While the architectural redundancy (RPR and ETR) ensures the existence of an optimal solution and the proprioceptive feedback provides the brain a mechanism to recognize and differentiate redundant muscular configurations, they do not imply how the optimal solution to the fundamental NP-optimization problem can be reached. We can only conjecture that the brain deploys a form of heuristic approximation approach to search for the near-optimal solution resembling our design strategy of the RS sensor. Similar approaches based on approximation, such as Inouye and Valero-Cuevas (2014) and Stanev and Moustakas (2018) have also been investigated for further understanding of the HMS, as well as for developing the control of anthropomorphic robots.

3.3 Deep Residual Networks. There exist several prominent uses of redundancy for enhancing classification accuracy in machine learning. In a “committee machine” or “ensemble learning,” multiple predictions are generated simultaneously by a collection of discrete instances that are based on the same or distinct predictive models. Because each instance produces a result with a different degree of error, an appropriate integration of these outcomes could lead to higher overall accuracy (Bishop, 2006). Another

approach is replication of individual neurons or subcircuits of an artificial neural network (ANN). Using mathematical models, Izui and Pentland (1990) Tanaka et al. (1988) conclude that replication can fundamentally alter the computation carried out by an ANN, resulting in quantitative enhancement of convergence speed, solution accuracy, and interconnection stability. The findings were used to design redundant ANNs simulating a robotic arm grasping an object in 2D space and a pattern-classification task with improved accuracy and convergence time (Medler & Dawson, 1994a, 1994b). These are prime examples of CRS, as shown in Figure 1b. Even though a marginal accuracy boost can be accomplished, the replication-based implementation prevents these systems from effectively employing redundancy without incurring excessive resource overhead.

Recently, deep learning has emerged as a leading field of machine learning (LeCun, Bengio, & Hinton, 2015). A feedforward deep neural network (DNN) produces a prediction by convoluting the inputs through various feature layers encoding the acquired knowledge (see Figure 5a). One of the breakthroughs in DNN design—the ResNet (He, Zhang, Ren, & Sun, 2016a, 2016b)—modifies the conventional structure by including “skip connections” or “identity mapping” that allow information to occasionally bypass an entire layer (see Figure 5b). Empirical experiments have demonstrated the superior predictive accuracy of ResNet compared to conventional networks with the same number of layers and parameters (He et al., 2016a, 2016b; Huang, Sun, Liu, Sedra, & Weinberger, 2016; Wu, Shen, & Hengel, 2016; Zagoruyko & Komodakis, 2017).

Although the advantage of ResNet is evident, many are baffled by how a subtle yet critical modification of the DNN could fundamentally alter its properties. It becomes clear as Veit et al. (2016) show that the ResNet’s behaviors resemble the characteristics of an ensemble of shallower networks. As illustrated in Figure 5c, the network can be “unraveled” as a sum of smaller subcircuits. Unlike the conventional DNN, where the input must be processed through all feature layers in a sequential order, information in a ResNet can flow through any one of the 2^N distinct pathways (N : number of layers) and is integrated only at the last step.

The structure of ResNet resembles that of an RES with both RPR and ETR properties. First, similar flows of information can now be accomplished by multiple different pathways of the network—the equivalence of microstates. Second, because of the entanglement among these pathways and microstates, an excessive level of information redundancy exponentially proportional to the number of layers is formulated without compromising the size of the network. In other words, the ResNet is a “redundant” network without being a “larger” network when compared to conventional DNN. The redundancy is embedded into the network’s architecture and is not simply an increase in the number of parameters. Although there is no immediate analog of the ResNet’s architecture to the biological counterparts, this example shows that redundancy with RPR and ETR properties

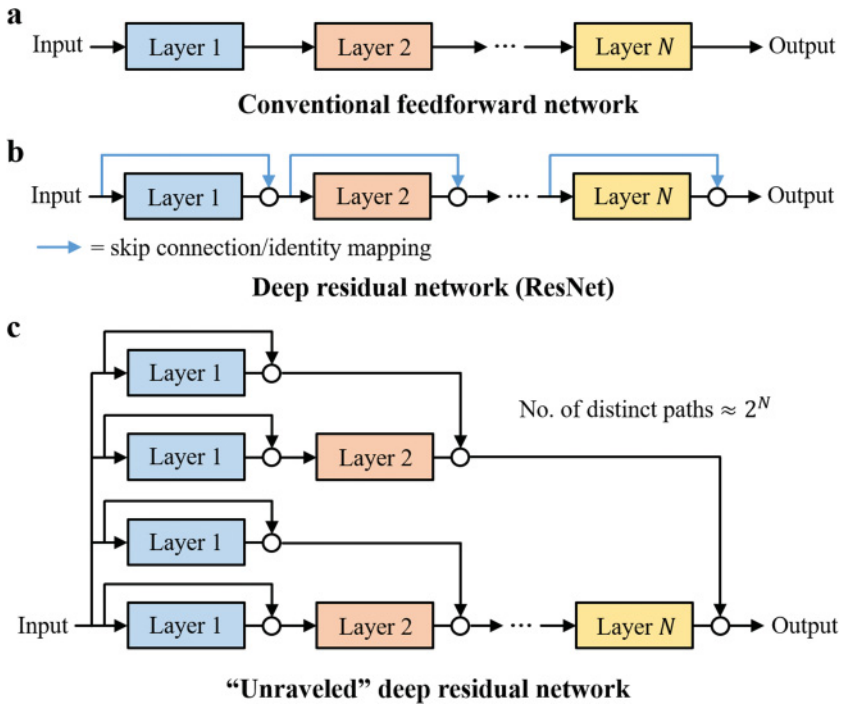


Figure 5: (a) A conventional feedforward DNN produces a prediction by convoluting the input through multiple feature layers. (b) A ResNet achieves superior predictive accuracy with the incorporation of skip connections or identity mappings, which allow information to bypass an entire layer. (c) It has been shown that the behavior of the ResNet is similar to a collection of shallower networks, resembling an RPR + ETR system (Veit, Wilber, & Belongie, 2016).

can be elegantly engineered into artificial systems, leading to major enhancement of performance.

4 Discussion and Conclusion

Although redundancy is no doubt an essential property of many biological processes, there are reasons to believe that its functions have not been fully appreciated, resulting in the absence in artificial designs. While the conventional interpretation often ties redundancy with fault tolerance, we propose a new model arguing that it can be engineered to advance the performance regarding accuracy and precision. Our theory highlights two fundamental mechanisms enabling such function: (1) RPR facilitates redundant encoding of information, and (2) ETR facilitates practical implementation of

information redundancy without physical redundancy. Besides suggesting the presence of these mechanisms in biological processes such as the human visual and musculoskeletal systems, we present two state-of-the-art man-made designs, the RS sensor architecture (Nguyen et al., 2016) and the ResNet (He et al., 2016a), where redundancy has been successfully employed. By providing new insights into these practical problems, we hope this letter will guide and motivate researchers in various fields of engineering and biological sciences to reexamine their interpretation of redundant systems and processes and come up with novel designs that incorporate redundancy in entirely different ways. In this sense, we believe our letter would have served its purpose and contributed great values to the scientific community.

Clearly, future work needs to be done to demonstrate the feasibility of such redundant architectures in practical application. First, under the guidelines of our framework, new engineering solutions should be derived to integrate redundancy into other designs for accuracy and precision enhancement. Although the principles of RPR and ETR are universal, their actual implementation varies drastically. The examples in electrical engineering and computer pointed out in this letter are merely the tip of the iceberg. Second, a new technique should be investigated to evaluate the information capacity of redundant systems that correlates to its upper bound of performance. A brute force approach used in the RS design (Nguyen et al., 2016) certainly cannot be applied to more complex systems such as ResNet. Finally, new methods should be developed to harness the full capacity of redundant systems. Redundant representation of information is irrelevant without an effective way to extract the optimal configuration. Almost all of the examples shown in this letter present NP-optimization problems for which solutions can be adequately obtained by means of approximation.

Appendix: Terminologies and Definitions

This section summaries new terminologies and definitions presented in this paper.

Conventional orthogonal system (COS): A nonredundant design where each entry of information is represented by a single unique system configuration (e.g., binary numeral system). It is in contrast to a redundant, nonorthogonal system.

Conventional redundant system (CRS): A redundant, nonorthogonal design achieved by replicating a partial or the entire system. In a CRS, information redundancy equals physical redundancy.

Entangled redundancy (ETR): The property of a redundant system where its microstates are partially correlated, allowing information redundancy to be implemented without excessive resource overhead or physical redundancy.

Entry of information: An enumerated symbol or state in an information processing system that may represent an input, output, or processing pathway.

Microstates: Distinct configurations of a redundant system that represent the same entry of information.

Redundant sensing (RS): A new design technique for using redundancy to enhance the precision of sensors and devices, in particular, analog-to-digital (AD) and digital-to-analog (DA) converters (Nguyen et al., 2016).

Representational redundancy (RPR): The property of a redundant, nonorthogonal system where each entry of information can be represented by numerous distinct configurations or microstates.

Representational and entangled redundant system (RES): A redundant design satisfying both RPR and ETR properties.

Acknowledgments

We thank the anonymous reviewers for insightful comments that helped improve the technical quality of this letter.

References

- Averbeck, B. B., Latham, P. E., & Pouget, A. (2006). Neural correlations, population coding and computation. *Nature Reviews Neuroscience*, 7(5), 358–366.
- Beck, J., & Schwartz, T. (1979). Vernier acuity with dot test objects. *Vision Research*, 19(3), 313–319.
- Bernstein, N. A. (1967). *The coordination and regulation of movements*. Oxford: Pergamon Press.
- Bishop, C. M. (2006). *Pattern recognition and machine learning*. New York: Springer.
- Cleather, D. J., & Bull, A. M. (2010). Lower-extremity musculoskeletal geometry affects the calculation of patellofemoral forces in vertical jumping and weightlifting. *Journal of Engineering in Medicine*, 224(9), 1073–1083.
- Forsberg, H., Eliasson, A., Kinoshita, H., Johansson, R., & Westling, G. (1991). Development of human precision grip I: Basic co-ordination of force. *Experimental Brain Research*, 85(2), 451–457.
- Glassman, R. B. (1987). An hypothesis about redundancy and reliability in the brains of higher species: Analogies with genes, internal organs, and engineering systems. *Neuroscience and Biobehavioral Reviews*, 11(3), 275–285.
- Hagen, D. A., & Valero-Cuevas, F. J. (2017). Similar movements are associated with drastically different muscle contraction velocities. *Journal of Biomechanics*, 59, 90–100.
- He, K., Zhang, X., Ren, S., & Sun, J. (2016a). Deep residual learning for image recognition. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (pp. 770–778). Piscataway, NJ: IEEE.

- He, K., Zhang, X., Ren, S., & Sun, J. (2016b). Identity mappings in deep residual networks. In *Proceedings of the European Conference on Computer Vision* (pp. 630–645). New York: Springer.
- Hennig, M. H., & Worgotter, F. (2004). Eye micro movements improve stimulus detection beyond the Nyquist limit in the peripheral retina. In S. Thrun, L. K. Saul, & B. Scholkopf (Eds.), *Advances in neural information processing systems* (pp. 1475–1482). Cambridge, MA: MIT Press.
- Hicheur, H., Zozor, S., Campagne, A., & Chauvin, A. (2013). Microsaccades are modulated by both attentional demands of a visual discrimination task and background noise. *Journal of Vision*, 13(13), 18–37.
- Huang, G., Sun, Y., Liu, Z., Sedra, D., & Weinberger, K. Q. (2016). Deep networks with stochastic depth. In *Proceedings of the European Conference on Computer Vision* (pp. 646–661). New York: Springer.
- Inouye, J. M., & Valero-Cuevas, F. J. (2014). Anthropomorphic tendon-driven robotic hands can exceed human grasping capabilities following optimization. *International Journal of Robotics Research*, 33(5), 694–705.
- Izui, Y., & Pentland, A. (1990). Analysis of neural networks with redundancy. *Neural Computation*, 2(2), 226–238.
- Kafri, R., Springer, M., & Pilpel, Y. (2009). Genetic redundancy: New tricks for old genes. *Cell*, 136(3), 389–392.
- Kutch, J. J., & Valero-Cuevas, F. J. (2011). Muscle redundancy does not imply robustness to muscle dysfunction. *Journal of Biomechanics*, 44(7), 1264–1270.
- LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature Neuroscience*, 521(7553), 436–444.
- Marjaninejad, A., & Valero-Cuevas, F. J. (2018). Should anthropomorphic systems be “redundant”? In G. Venture, J.-p. Laumond, & B. Watier (Eds.), *Biomechanics of Anthropomorphic Systems* (pp. 7–34). New York: Springer.
- Martinez-Conde, S., Otero-Millan, J., & Macknik, S. L. (2013). The impact of microsaccades on vision: Towards a unified theory of saccadic function. *Nature Reviews Neuroscience*, 14(2), 83–96.
- Medler, D. A., & Dawson, M. R. W. (1994a). Training redundant artificial neural networks: Imposing biology on technology. *Psychological Research*, 57(1), 54–62.
- Medler, D. A., & Dawson, M. R. W. (1994b). Using redundancy to improve the performance of artificial neural networks. In *Proceedings of the Biennial Conference of the Canadian Society for Computational Studies of Intelligence on Advances in Artificial Intelligence* (pp. 131–138).
- Moissenet, F., Cheze, L., & Dumas, R. (2016). Influence of the level of muscular redundancy on the validity of a musculoskeletal model. *Journal of Biomechanical Engineering*, 138(2), 021019.
- Narayanan, N. S., Kimchi, E. Y., & Laubach, M. (2005). Redundancy and synergy of neuronal ensembles in motor cortex. *Journal of Neuroscience*, 25(17), 4207–4216.
- Nguyen, A. T., Xu, J., & Yang, Z. (2015). A 14-bit 0.17mm² SAR ADC in 0.13μm CMOS for high precision nerve recording. In *Proceedings of the IEEE Custom Integrated Circuits Conference* (pp. 1–4). Piscataway, NJ: IEEE.
- Nguyen, A. T., Xu, J., & Yang, Z. (2016). A bio-inspired redundant sensing architecture. In D. D. Lee, M. Sugiyama, U. V. Luxburg, I. Guyon, & R. Garnett (Eds.), *Advances in neural information processing systems*, 29 (pp. 2379–2387).

- Nowak, M. A., Boerlijst, M. C., Cooke, J., & Smith, J. M. (1997). Evolution of genetic redundancy. *Nature Neuroscience*, 388(6638), 167–171.
- Panzeri, S., Schultz, S. R., Treves, A., & Rolls, E. T. (1999). Correlations and the encoding of information in the nervous system. *Proceedings of the Royal Society of London B: Biological Sciences*, 266(1423), 1001–1012.
- Pitkow, X., & Angelaki, D. E. (2017). Inference in the brain: Statistics flowing in redundant population codes. *Neuron*, 94(5), 943–953.
- Schreuders, T. A., Selles, R. W., Roebroek, M. E., & Stam, H. J. (2006). Strength measurements of the intrinsic hand muscles: A review of the development and evaluation of the rotterdam intrinsic hand myometer. *Journal of Hand Therapy*, 19(4), 393–402.
- Shannon, C. E. (1948). A mathematical theory of communication. *Bell System Technical Journal*, 27, 379–423, 623–656.
- Stanev, D., & Moustakas, K. (2018). Simulation of constrained musculoskeletal systems in task space. *IEEE Transactions on Biomedical Engineering*, 65(2), 307–318.
- Tanaka, H., Matsuda, S., Ogi, H., Izui, Y., Taoka, H., & Sakaguchi, T. (1988). Redundant coding for fault tolerant computing on Hopfield network. *Neural Networks*, 1, 141.
- Tautz, D. (1992). Redundancies, development and the flow of information. *Bioessays*, 14(4), 263–266.
- Valero-Cuevas, F. J. (2015). *Fundamentals of neuromechanics*. New York: Springer.
- Valero-Cuevas, F. J., Cohn, B. A., Yngvason, H. F., & Lawrence, E. L. (2015). Exploring the high-dimensional structure of muscle redundancy via subject-specific and generic musculoskeletal models. *Journal of Biomechanics*, 48(11), 2887–2896.
- Veit, A., Wilber, M., & Belongie, S. (2016). Residual networks behave like ensembles of relatively shallow networks. In D. D. Lee, M. Sugiyama, U. V. Luxburg, I. Guyon, & R. Garnett (Eds.), *Advances in neural information processing systems* (pp. 550–558). Red Hook, NY: Curran.
- Wu, Z., Shen, C., & Hengel, A. V. D. (2016). *Wider or deeper: Revisiting the resnet model for visual recognition*. arXiv:1611.10080.
- York, R. A., & Fernald, R. D. (2017). The repeated evolution of behavior. *Frontiers in Ecology and Evolution*, 4(143), 1–10.
- Zagoruyko, S., & Komodakis, N. (2017). Wide residual networks. arXiv:1605.07146.