Asymmetric Change-of-Probability Measures for Tail Risk Management

Saeede Enayati
Electrical and Computer Engineering
University of Massachusetts, Amherst
senayati@umass.edu

Hossein Pishro-Nik
Electrical and Computer Engineering
University of Massachusetts, Amherst
pishro@umass.edu

***Abstract*—Large (extreme, high-impact) events can occur in complex systems as a result of fat-tailed distributions of the system behavior. Decision-making in the presence of a potential high-impact event where the law of large numbers (LLN) cannot be applied, is not as straightforward as decision-making in a normal scenario where the LLN is used. In this paper, a general framework is introduced for decision-making in the presence of a potential high-impact event using change-of-probability measures. The idea of the proposed framework is to weigh the high negative consequences of non-LLN decision-making problems as a tail risk management strategy. The proposed approach is named asymmetric change-of-probability measures (ACM) as the right and the left tails of the distributions are treated asymmetrically. A key to this approach is to define and satisfy required properties so that the change-of-measure operation is performed in a principled way. An important property is ensuring upper bounds for the relative entropy between the distributions. We first introduce asymmetric bounded expectation (ABE), as a special case of the general approach. We then extend the proposed asymmetric method to the general change-of-measure. Benefiting from the same properties as the symmetric change-of-measure, we show that the asymmetric approach can be potentially a promising method for decision-making under non-LLN risk management circumstances in complex systems. Through a practical example from venture capital (VC) in finance, and in comparison to the symmetric change-of-measure, we show that considering tail risk management will result in a different decision-making outcome where the VC is required to invest in more startups to avoid a loss.**

***Index terms*—** Decision-making, complex systems, risk management, expected utility, relative entropy, change-of-measure.

I. Introduction

Financial markets, Internet, electrical power network, transportation systems, spread of contagious diseases etc, are only a handful of the popular complex systems among many of them [1]–[3]. The presence of power-law in complex systems has been recognized to describe many behaviors of such systems [4]–[6]. Besides, the interconnectedness of a complex system and the interaction between the physical parts and human or intelligent agents demands frequent decision-makings for each. Decisions that are made based on probabilities and their corresponding results, known as probabilistic decision-making, and failing in which will produce *large events* [7] or the so-called high-impact events.

Probabilistic decision-making has been investigated in many different disciplines such as information theory, artificial intelligence, psychology, economics, and other social sciences [8]–[16]. It is important to have a solid framework for decision making, specifically, as more decisions will be made by artificial intelligence (AI) agents in near future, [17], [18] and when there are high-impact decisions with high undesired consequences and very small probabilities [19].

A very common approach to decision making is the principle of maximizing expected utility. While very useful, the notion of expected utility and probability measures have a fundamental limitation: They are valid when the law of large numbers (LLN) holds, i.e., a long run of the experiment is available. This is not the case usually when we deal with complex systems where we face non-LLN decision-making problems. Hence, there seems to be a lack of rigorous framework to deal with decision-making in such contexts.

In our recent work, [20], we proposed a new decision-making framework based on change-of-probability measures to provide a systematic approach for rational decision-making under non-LLN regimes. The main idea was to apply a change-of-measure operation to amplify the more likely outcomes and weaken the less likely ones, which indeed makes more sense in a single run (or a few runs) of an experiment. A main contribution of [20] was to ensure such change-of-measure operations are done in a principled way. The proposed method was shown to be a generalization of the expected utility theory (EUT) from several perspectives.

Nevertheless, in decision-making for complex systems, a very important goal is to manage risks. In our framework, this is associated to left-tail risk, i.e., when the value of utility is the least. While the proposed framework in [20] can account for accumulation of risk, it is nevertheless symmetric in the sense that it treats left tails and right tails in the same manner. This is not usually sufficient to handle left-tail risks [21].

The contributions of this paper are as follows:

- We develop a more general change-of-probability measures framework that is capable of handling risk management strategies in complex systems. The idea here is to ensure that left-tails are sufficiently included in our decision making framework. The method is referred to as asymmetric change-of-probability measures (ACM). This is achieved by introducing the parameter k that controls for risk aversion level.
- We provide carefully chosen properties that need to be satisfied by such change-of-measure operations while at the same time ensuring that the relative entropy between the resulting distributions are sufficiently bounded.

This work was supported by NSF under grants CNS-1932326 and CNS-2150832.

- We prove that the proposed method converges to expected utility if the number of experiments grows.
- We provide a general method to achieve such change-of-measure operations (using consistent functions).
- Upper bounds on the distortion parameter (relative entropy) are obtained. In fact, another advantage of the proposed method is the use of relative entropy as opposed to total variation distance in [20]. This is specially important as it is suitable for future work, where communication and control costs are incorporated in the decision making [22], [23].
- Finally, through a practical example from finance, we show how the ACM will affect the decision made by a venture capital (VC) in comparison to his decision using the symmetric change-of-measure.

A. Related Works

Decision-making is a multi-disciplinary problem from information theory and engineering systems to medicine and social sciences applications [8]–[16], [24]–[27].

There have been two main attitudes towards decision-making problems: descriptive theories and normative theories.

In descriptive theory, the focus is on analyzing how human makes decisions from a psychological point of view [28]. On the other hand, in normative works, the goal is to design efficient methodologies in order to make a rational decision [29]. The most common approach in order to make a rational decision suggested by normative theories, is the principal of maximum expected utility which is widely used in different disciplines [30].

The limitations of expected utility approach have been discussed extensively, for example, in the contexts of problems such as St. Petersburg paradox [31]. In [20], we unified such limitations by formulating them under non-LLN regimes. This allowed us to provide a framework that converged to expected utility as the number of repetitions grow; while at the same time, it delivered reasonable results when the number of repetitions was small. In this paper, with a focus on risk management, we generalize the work in [20] to allow for risk management. Change-of-probability measures has been exploited in different contexts before. For example in information theory [32]; signal processing [33], [34]; and fuzzy measure theory [35].

B. Organization

This paper is organized as follows: In Section II we introduce the asymmetric bounded expectation (ABE) as well as upper bounds for relative entropy. Section III provides the general ACM. Section IV provides the systematic approach and finally, Section V concludes the paper.

As this paper is a generalization of the work in [20], some of the needed proofs can be easily obtained by adapting the proofs in [20]. Such proofs are eliminated for brevity and lack of space.

II. Asymmetric Bounded Expectation

We start by introducing a very simple ACM operation: ABE. Due to its simplicity, ABE is very helpful in obtaining insights about the general ACM operation. Consider a scenario where an agent is considering one of m possible actions or choices. The random variables that represent the rewards (utilities) of potential actions are represented as X_i, for $i = 1, 2, \cdots, m$. For clarity of exposition and to avoid measure theoretical technicalities, we assume the random variables X_i are discrete and the countable set R_X includes all the possible values for these random variables. Note that all the arguments can be extended to general random variables similar to what has been done [20].

Basically, bounded expectation (BE) introduced in [20] has a very intuitive and *interpretable* definition. It is motivated by the de minimis risk principle [36] which states that we should ignore very small probabilities, say below ϵ. Hence, the basic idea of the BE is very simple: We first identify "extreme values" (outliers) of random variables $X_i, i = 1, 2, \ldots, m$ from the right and left in such a way that the probabilities of such extreme values are in total less than or equal to $\frac{\epsilon}{m}$.

ABE is further the modified version of BE. The basic idea is pretty similar: We first identify outliers of X_i from the left and right in such a way that the probabilities of such extreme values are in total less than or equal to $\frac{\epsilon}{m}$ where $\epsilon = \epsilon_l + \epsilon_r$ is what we consider the rationally negligible probability. The ABE of X_i, shown as $E_\epsilon[X_i]$, is then the conditional expected value of X_i given that X_i is not in the outlier region.

Hence, in symmetric BE we removed a mass of $\frac{\epsilon}{2}$ unlikely events from both tails of the distribution while in asymmetric BE, we choose different thresholds ϵ_l (for the left tail) and ϵ_r (for the right tail). Specifically, to focus on the risk, we ensure $\epsilon_r > \epsilon_l$. To easily apply our systematic approach, we could require $\epsilon_r = k\epsilon_l$, where $k > 1$. The parameter k shows the level of risk aversion. An extremely risk averse agent might set a very large value for k, i.e., as $k \to \infty$, then $\epsilon_l \to 0$ and $\epsilon_r \to \epsilon$.

To have a general definition that can be applied to all kinds of random variables, we define ABE as follows. We can generate N random sample of $X_{(i)}$s (independently), sort the values increasingly and throw out the outliers. Doing so, the index set of the remaining samples, known as "normal set" will be obtained as below

$$I_N = \{\lfloor \frac{N\epsilon_l}{m} \rfloor, \lfloor \frac{N\epsilon_l}{m} \rfloor + 1, \ldots, N - \lfloor \frac{N\epsilon_r}{m} \rfloor\}, \qquad (1)$$

and the outlier index set will be $\{1, 2, \ldots, N\} - I_N$. ABE is then the sample mean of the values in I_N as below:

$$E_\epsilon[X] = \lim_{N\to\infty} \frac{1}{|I_N|} \sum_{i\in I_N} X_{(i)}, \qquad (2)$$

It is not difficult to see that the limit exists and is finite using the LLN.

Let us denote the primary probabilities of a discrete random variable X by $P = \{p_1, p_2, \ldots\}$ (i.e., $p_i = P(X = x_i)$) and denote the corresponding changed probability values by $Q = \{q_1, q_2, \ldots\}$. Also, let $D(Q||P)$ show the relative entropy between Q and P. In the following lemma, an upper bound for $D(Q||P)$ can be obtained:

Lemma 1. *If Q is absolutely continuous with respect to P, $Q \ll P$, the relative entropy of Q and P is upper bounded as*

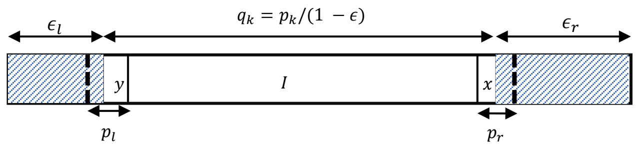

Fig. 1. Partitioning of the probabilities and their corresponding changed probability measures. According to the ABE, the probability of the samples in the shaded region will be changed to 0.

below:

$$D(Q||P) \leq \frac{1}{\ln 2} \max_k \left(\frac{q_k}{p_k}\right) |P-Q|, \tag{3}$$

where $|P-Q|$ is the total variation distance between P and Q.

Proof.

$$\begin{aligned} D(Q||P) &= \sum_k q_k \log_2 \frac{q_k}{p_k} \\ &\leq \frac{1}{\ln 2} \sum_k q_k \left(\frac{q_k}{p_k} - 1\right) \\ &= \frac{1}{\ln 2} \sum_{q_k > p_k} \frac{q_k}{p_k} (q_k - p_k) + \frac{1}{\ln 2} \sum_{q_k < p_k} \frac{q_k}{p_k} (q_k - p_k) \\ &\leq \frac{1}{\ln 2} \sum_{q_k > p_k} \frac{q_k}{p_k} (q_k - p_k) \\ &\leq \frac{1}{\ln 2} \max_k \frac{q_k}{p_k} \sum_{q_k > p_k} (q_k - p_k) \\ &= \frac{1}{\ln 2} \max_k \frac{q_k}{p_k} |P-Q|, \end{aligned}$$

where the first inequality results from the inequality $\ln x \leq x - 1$, for $x > 0$. ■

As Lemma 1 is a general axiom, we can use it to obtain an upper bound for the relative entropy between the distributions in the case of ABE.

Corollary 1. *For ABE, the upper bound for $D(Q||P)$ can be obtained as below:*

$$D(Q||P) \leq \frac{1}{\ln 2} \frac{\epsilon}{1-\epsilon},$$

since, in ABE, $\max_k \frac{q_k}{p_k} = \frac{1}{1-\epsilon}$ and $|P-Q| = \epsilon$.

Nevertheless, for the special case of ABE, we can further obtain a tighter bound through the next lemma.

Lemma 2. *The relative entropy of Q and P under ABE is upper bounded as below:*

$$D(Q||P) \leq \log_2 \frac{1}{1-\epsilon}. \tag{4}$$

Proof. After change-of-measure by ABE, we partition the probability space as in Figure 1.

Now, we calculate the relative entropy:

$$\begin{aligned} D(Q||P) &= \sum_k q_k \log_2 \frac{q_k}{p_k} \\ &= \sum_{k \in I} \frac{p_k}{1-\epsilon} \log_2 \frac{1}{1-\epsilon} + \frac{x}{1-\epsilon} \log_2 \frac{x}{(1-\epsilon)p_r} \\ &+ \frac{y}{1-\epsilon} \log_2 \frac{y}{(1-\epsilon)p_l} + 0 \\ &= \frac{1}{1-\epsilon} \log_2 \frac{1}{1-\epsilon} (S + x + y) \\ &+ \frac{x}{1-\epsilon} \log_2 \frac{x}{p_r} + \frac{y}{1-\epsilon} \log_2 \frac{y}{p_l} \\ &\leq \log_2 \frac{1}{1-\epsilon}, \end{aligned}$$

where $S = \sum_{k \in I} p_k$, and the last inequality comes from the fact that the last two terms are negative, since $x < p_r$ and $y < p_l$, and $S + x + y = 1 - \epsilon$. ■

Lemma 2 obtains a tighter upper bound for $D(Q||P)$ (for the special case of ABE) than Corollary 1 as

$$D(Q||P) \leq \log_2 \frac{1}{1-\epsilon} \leq \frac{1}{\ln 2} \frac{\epsilon}{1-\epsilon}, \quad \text{for } \epsilon > 0.$$

Although ABE benefits from a simple and intuitive nature, it has limitations too. The main limitation is that it partitions the probability measures into two sets: normal and outliers sets. Hence, it causes a sharp change in probability measures. In the next section, we develop a framework that not only partitions the probability space in to more than two parts, but also changes the probability measures more smoothly.

III. General Asymmetric Change-of-Probability Measures

Similar to the symmetric change-of-measure, in order to be a principled operation, ACM needs to meet several important properties. These properties are fairly similar to that of symmetric change-of-measure developed in [20]. In the following, we briefly revisit them along with the mentioning the possible differences with the symmetric case.

A. Asymmetric ϵ_d-Consistent Change-of-Measure Policies

We consider a complete probability space $(\Omega, \mathcal{F}, P)$ where the random variables $X_i : \Omega \mapsto \mathbb{R}$, for $i = 1, 2, \cdots, m$ represent the utilities. It is in general convenient (and not restrictive) if we assume $\mathcal{F}$ is the sigma field generated by all the involved random variables.

The goal here is to define a new probability measures Q_{X_i} on $(\Omega, \mathcal{F})$ to be used in evaluating the true value of these actions. Specifically, we would construct mapping

$$\{X_1, X_2, \cdots, X_m, P\} \overset{ch}{\mapsto} \{Q_{X_i}\},$$

such that for any $X \in \{X_1, X_2, \cdots, X_m\}$, its value $v[X]$ is given by

$$v[X] = v_X[X] = \int_\Omega X(\omega) dQ_X(\omega).$$

We might require that all Q_{X_i} are the same, $Q_{X_i} = Q$. In this case, we call such change-of-measure operation uniform. For simplicity of notation, we assume uniformity in this section.

Note that although the idea of change-of-measure is to amplify the most likely outcomes while weakening the highly unlikely ones, by considering ACM, we wish to take the probability of large risks in to account so that we have a robust decision-making policy. Furthermore, our goal is to describe mappings $P \mapsto Q$ that have desirable properties consistent with probabilistic decision-making.

In order to be a valid operation, change-of-measure needs to satisfy some properties. There is a lot of similarity to the properties introduced in [20] for symmetric case. The main difference is in Property 5 and Property 6. For the sake of readers convenience, we revisit the properties without mentioning the details and proofs.

Property 1. *(Finiteness)*

1) *If $P(\{\omega \in \Omega : X(\omega) < \infty\}) = 1$, then $v[X] < \infty$.*
2) *If $P(\{\omega \in \Omega : X(\omega) > -\infty\}) = 1$, then $v[X] > -\infty$.*

Property 2. *(Dominance)*
If $P(\{\omega \in \Omega : X_1(\omega) \leq X_2(\omega)\}) = 1$, then $v[X_1] \leq v[X_2]$.

Property 3. *(CDF Sufficiency/Symmetry) Consider m random variables $X_i : \Omega \mapsto \mathbb{R}$ for $i = 1, 2, \cdots, m$ with the joint cumulative distribution function (CDF) $F_{X_1,X_2,...,X_m}(x_1, x_2, ..., x_m)$, and let $[X_1, X_2, \cdots, X_m] \mapsto [v_1, v_2, \cdots, v_m]$.*

1) *The values $v_i = v[X_i]$, for $i = 1, 2, \cdots, m$, are uniquely determined by $F_{X_1,X_2,...,X_m}$.*
2) *For any permutation $\pi : \{1, 2, \cdots, m\} \mapsto \{1, 2, \cdots, m\}$, we must have*

$$[X_{\pi(1)}, X_{\pi(2)}, \cdots, X_{\pi(m)}] \mapsto [v_{\pi(1)}, v_{\pi(2)}, \cdots, v_{\pi(1)}].$$

Property 4. *(Convergence) For a sequence of random variables $X_1^{[n]}$, $n = 1, 2, 3, \cdots$, on $(\Omega, \mathcal{F}, P)$, where $Q^{[n]}$ is the corresponding measure and*

$$X_1^{[n]} \xrightarrow{a.s.} X_1,$$

where " $\xrightarrow{a.s.}$ " indicates almost sure convergence (with respect to P), and all the $X_1^{[n]}$s are dominated in absolute value by an integrable (with respect to P and $Q^{[n]}$) random variable Y, we have

$$\lim_{n\to\infty} v^{[n]}[X_1^{[n]}] = v[X_1].$$

Note that since $Q^{[n]} \ll P$, almost sure convergence with respect to P also ensures almost sure convergence with respect to all $Q^{[n]}$s.

Property 5. *(Weak linearity) $v[aX + b] = av[X] + b$, for any $a \geq 0, b \in \mathbb{R}$.*

Note that unlike the symmetric case where this property holds for all $a \in \mathbb{R}$ [20], for the ACM, Property 5 holds only for $a \geq 0$.

Property 6. *(Bounded relative entropy distortion) For any event $B \in \sigma(X_1, X_2, \cdots, X_m)$, we must have*

$$D(Q||P) \leq \epsilon_d.$$

Now, we have the ϵ_d-consistent policy definition:

Definition 1. *We say that a change-of-measure operation is an asymmetric ϵ_d-*consistent *policy if it satisfies properties 1 through 6.*

If a decision is being repeated several times independently and we define

$$\overline{X}_n = \frac{X^{(1)} + X^{(2)} + ... + X^{(n)}}{n},$$

then $v[\overline{X}_n]$ converges to $E[X]$.

Theorem 1. *(Limit Theorem) Let $X^{(1)}$,$X^{(2)}$,..., $X^{(n)}$ be independent and identically distributed (i.i.d.) random variables with expected values $E[X^{(i)}] = \mu < \infty$. Let $v[.]$ be associated with an asymmetric ϵ_d-consistent change-of-measure policy. Assume $\overline{X}_n$s are dominated in absolute value by an integrable random variable Y. We have*

$$\lim_{n\to\infty} v[\overline{X}_n] = \mu.$$

Note that the above theorem essentially guarantees that the proposed method converges to the expected utility as n becomes large. Nevertheless, the value of the proposed approach lies in the case where we "cannot" repeat the decision-making process a large number of times. As this is the case in almost all high-impact decisions, the proposed method can potentially be a promising approach. Besides, the asymmetric nature of the operation, makes it well-suited for risk management.

IV. Systematic Approach to Constructing Asymmetric ϵ_d-Consistent Policies

In this section, we first introduce *consistent functions* for ACM and then we develop the systematic procedure to use the proposed method in decision-making.

Definition 2. *(Consistent functions) We say that a function $g : [0, 1] \mapsto [0, 1]$ is a* consistent *function with respect to random variables $X_1, X_2, \cdots, X_m$ if all of the following conditions are satisfied:*

1) *g is continuous and increasing, and $g(0) = 0, g(1) = 1$.*
2) *(Lipschitz continuity) There exists $c_g \in \mathbb{R}$ such that $|g(x)-g(y)| \leq c_g|x-y|$ for all $x \in [0, 1]$ and $y \in [0, 1]$.*
3) *On the interval $\left[0, \frac{1}{2}\right]$, $g(\cdot)$ is convex and $g(x) \leq x$.*
4) *On the interval $\left[\frac{1}{2}, 1\right]$, $g(\cdot)$ is concave and $g(x) \geq x$.*
5) *For some $k \in \mathbb{N}$ (risk aversion parameter), we have*
$$g(x) + kg(1 - x) + (k - 1)x = k, \qquad \text{for all } x \in [0, .5].$$
6) *For each $i = 1, 2, \cdots, m$, there are constants c_i and c_i' in $\mathbb{R}$ such that*
$$\int_{c_i}^{\infty} g(\bar{F}_i(x))dx < \infty, \quad \int_{-\infty}^{c_i'} g(F_i(x))dx < \infty.$$

Now, we have the following theorem:

Theorem 2. *Let $g : [0,1] \mapsto [0,1]$ be a consistent function with respect to random variables $X_1, X_2, \cdots, X_m$. For any $i \in \{1,2,\cdots,m\}$ and $x \in \mathbb{R}$, define*

$$Q_{X_i}(X_i > x) = g(P(X_i > x)).$$

Then, $P \mapsto Q_X$ is a nonuniform asymmetric ϵ_d-consistent change-of-measure policy, where

$$\epsilon_d \leq \frac{2c_g}{\ln 2} \sup_{x\in[0,1]} |g(x) - x|. \tag{5}$$

Proof. (Proof of Theorem 2) Properties 1-4 can be proved similar to the symmetric change-of-measure as in [20]. Also, Property 5 only holds for $a \geq 0$. Finally, to prove Equation (5), remember that according to the Lemma 1, we have

$$D(Q||P) \leq \frac{1}{\ln 2} \max_k \left(\frac{q_k}{p_k}\right) |P - Q|.$$

Hence, we need to calculate the term $\max_k \frac{q_k}{p_k}$ based on the consistent function. Therefore, using Equation (7), we write it as below:

$$\frac{q_k}{p_k} = \frac{g\left(\sum_{i=k}^{\infty} p_i\right) - g\left(\sum_{i=k+1}^{\infty} p_i\right)}{p_k} \leq c_g,$$

where the inequality comes from the Lipschitz continuity in Property 2 of Definition 2. Also, noting that [20]

$$|Q - P| \leq 2 \sup_{x\in[0,1]} |g(x) - x|,$$

we obtain:

$$D(Q||P) \leq \frac{2c_g}{\ln 2} \sup_{x\in[0,1]} |g(x) - x|. \tag{6}$$

■

For example, for $m = 1$ and non-atomic probability spaces, $g(x)$ of the form below results in ABE

$$g(x) = \begin{cases} 0 & x < \epsilon_r \\ \frac{x-\epsilon_r}{1-2\epsilon_r} & \epsilon_r \leq x \leq \frac{1}{2} \\ \frac{x-\epsilon_l}{1-2\epsilon_l} & \frac{1}{2} \leq x \leq 1-\epsilon_l \\ 1 & x > 1-\epsilon_l \end{cases}.$$

A. A Systematic Approach to Decision-Making in Non-LLN Regimes

Here, we provide specific constructions of consistent functions, and briefly discuss how it can be used in a systematic approach for decision-making in non-LLN regimes using ACMs. In fact it can be developed similar to the symmetric case. The only difference is the consistent function itself which is no longer symmetric. It is instead asymmetric in such a way that it distorts the left tail less than the right tail. Other steps of the process is the same as before. For the reader's convenience, we revisit the process developed in [20] in the sequel.

We start by picking an ϵ-consistent policy such as the method described in the previous section using consistent functions. We note that for the ACM, for a given k, we have $\epsilon = (k+1)\epsilon_l$ and if we let $\epsilon = 0$, we obtain the same results derived from expected utility theory. In general, let $i(\epsilon)$ be the preferred option for a specific ϵ. As we then increase ϵ, we take note of the possible changes in $i(\epsilon)$. Let ϵ^* be the value of ϵ where the first change occurs in $i(\epsilon)$, i.e., $i(\epsilon^*) \neq i(0)$. The key insights are as follows:

1) The larger the value of ϵ^*, the more stable is the choice made by the expected utility $(i(0))$. That is, it is more likely that the expected utility is suggesting a good option.
2) On the other hand, if the value of ϵ^* is small, this is a high indication that $i(\epsilon^*)$ might be the best choice.

As the problem of decision-making under non-LLN regimes is multifaceted and most likely a simple narrow approach will not be enough, the proposed method above, where we look at how the preferences change as ϵ changes, seems to be a step in the right direction. An interesting question for further research seems to be finding guidelines on the choice of the threshold value of ϵ^* at which $i(\epsilon^*)$ becomes the preferred option. As a very rough rule of thumb, one might suggest $\epsilon^* < 0.05$ might be used as the threshold.

Here is an example of a consistent function:

$$g(x) = \begin{cases} 0 & x \leq \epsilon_r \\ 2^{\alpha} x^{1+\alpha} & \epsilon_r \leq x \leq 0.5 \\ \frac{1+(k-1)x-g(1-x)}{k} & 0.5 \leq x \leq 1-\epsilon_l \\ 1 & x \geq 1-\epsilon_l \end{cases},$$

where α is a parameter that determines the distortion parameter, ϵ_d. For $\alpha = 0$ we have $g(x) = x$, so we obtain the standard expected utility, and $\epsilon_d = 0$. Furthermore, k determines the level of risk aversion. Figure 2 shows the asymmetric $g(x)$ for different α and k values.

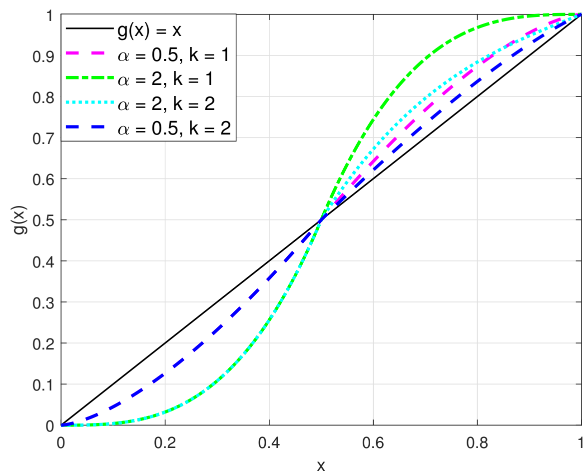

Fig. 2. A representation of $g(x)$ with different values of α and k.

Having a consistent function, we can then use the systematic approach proposed in [20]. The main difference here is that we are using an asymmetric change-of-measure and this will allow us to control for the level of acceptable risk.

Specifically, if X is a discrete random variable and bounded from the left, we can simplify the change-of-measure operation in Theorem 2 in the following way. Suppose $\{x_1, x_2, x_3, \cdots\}$ are potential values of X in an ordered way, i.e.,

$$x_1 < x_2 < x_3 \cdots.$$

Let $p_i = P(X = x_i)$. Then, we obtain the changed probabilities, q_j, for $j = 1, 2, \ldots,$ as below

$$q_j \triangleq Q_i(X = x_j) = g\left(\sum_{i=j}^{\infty} p_i\right) - g\left(\sum_{i=j+1}^{\infty} p_i\right). \quad (7)$$

If the range is finite, i.e.,

$$x_1 < x_2 < x_3 \cdots < x_r,$$

then, for $j = 1, 2, \cdots, r-1$, we obtain

$$q_j = g\left(\sum_{i=j}^{r} p_i\right) - g\left(\sum_{i=j+1}^{r} p_i\right),$$

and

$$q_r \triangleq Q_i(X = x_r) = g(p_r).$$

The value of ϵ_d can be obtained by calculating the relative entropy between P and Q:

$$\epsilon_d = \sum_{i=1}^{r} q_i \log_2 \frac{q_i}{p_i}.$$

The value of X is then obtained as

$$v[X] = v_\epsilon[X] = \sum_j x_j q_j.$$

B. Example from Financial Market

Here, we revisit the problem of angel and venture capital (VC) investment from [20] and apply the ACM systematic approach. The problem is as follows:

An angel or a venture capital investment fund is supposed to invest in technology startups. A fundamental question is in how many startups the available funds should be divided with a risk minimization approach? To answer this question, we apply the proposed change-of-measure-based approach in [20] and the ACM developed in this paper and compare their results.

We use the same formulation as in [20]. Hence, we denote by L the total number of companies in which the fund invests. Also, $X^{(j)}$, $j = 1, 2, \cdots, L$ denotes the total profit from the investment in the jth company assuming one unit of money being invested. For example, if the j^{th} company fails, we let $X^{(j)} = -1$ and if the investor triples the invested amount, we let $X^{(j)} = 2$. For simplicity, we assume that the fund invests equal amounts in each company. Furthermore, a typical distribution of $X^{(j)}$s is shown in Figure 3.

Now, assume that the investor requires a minimum profit of 140% over the length of the investment, which is usually a few years for each startup. To reach to the desired profit, and considering a 0.05 tolerance (i.e., $\epsilon^* = 0.05$), we have seen in [20] that with a symmetric change-of-measure, where $k = 1$, the VC needs to invest in at least 12 startups.

However, by applying the tail risk management approach we get different results. For example, for $k = 2$, we obtain $L \geq 19$ and for $k = 5$, we obtain $L \geq 30$. In other words, in order to avoid the risk, the VC is required to apply in more startups. The results are shown in Table for readers convenience.

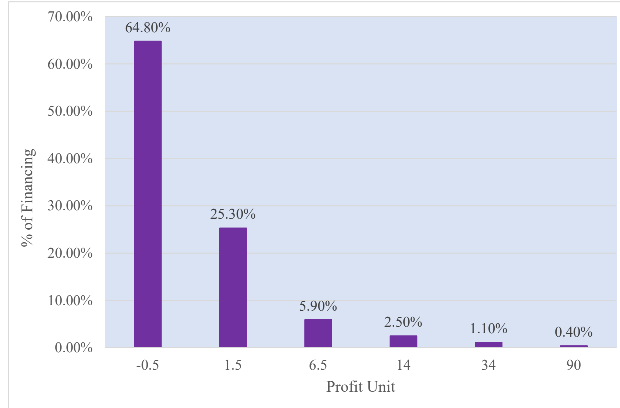

Fig. 3. Distribution of U.S. venture returns between 2004 and 2013, adapted from [37].

TABLE I
THE VC PROBLEM SOLUTION FOR DIFFERENT APPROACHES.

$k = 1$	$k = 2$	$k = 5$
$L \geq 12$	$L \geq 19$	$L \geq 30$

V. CONCLUSION

Complex systems potentially suffer from occurrence of high-impact events with tragic consequences. In this paper, we introduced asymmetric change-of-probability measures (ACM) as a method to deal with the risk management problems in the complex systems. Considering the non-LLN nature of the problem, the goal was to ensure that the left tails are sufficiently included in the decision-making process. We first applied the ACM to the bounded expectation as a special case, and obtained the upper bounds for the relative entropy. We then developed a general theory for ACM and discussed that it benefits from most of the symmetric change-of-measure properties. Hence, it can be a principled methodology for risk management under the non-LLN regimes in complex systems. Finally, we revisited the VC problem and solved it by the proposed risk management approach. We showed that the VC needs to increase the number of startups for his investment so that the risk probability is kept as low as desired.

REFERENCES

[1] M. C. Leles, E. F. Sbruzzi, J. M. de Oliveira, and C. L. Nascimento, "A MatLab computational framework for multiagent system simulation of financial markets," in *IEEE International Systems Conference (SysCon)*, Orlando, FL, USA, April 2019, pp. 1–8.

[2] F. Shirvani, W. Scott, G. Kennedy, F. Rezaeibagha, and P. Campbell, "Developing a modelling framework for aligning the human aspects to the physical system in large complex systems," in *IEEE International Systems Conference (SysCon)*, Orlando, FL, USA, 2019, pp. 1–6.

[3] P. Fieguth, *An Introduction to Complex Systems-Society, Ecology, and Nonlinear Dynamics*. Springer, Switzerland, 2017.

[4] X. Gabaix, P. Gopikrishnan, V. Plerou, and H. E. Stanley, "A theory of power-law distributions in financial market fluctuations," *Nature*, vol. 423, no. 6937, pp. 267–270, May 2003.

[5] A. M. Lopes and J. A. T. Machado, "Power law behaviour in complex systems," *Entropy*, vol. 20, no. 9, 2018. [Online]. Available: https://www.mdpi.com/1099-4300/20/9/671

[6] G. L. Vasconcelos, A. M. Macêdo, G. C. Duarte-Filho, A. A. Brum, R. Ospina, and F. A. Almeida, "Power law behaviour in the saturation regime of fatality curves of the COVID-19 pandemic," *Scientific Reports*, vol. 11, no. 1, pp. 1–12, Feb. 2021.

[7] S. E. Page, *Understanding complexity*. Teaching Company Chantilly, VA, 2009.

[8] A. Cohen, N. Merhav, and T. Weissman, "Scanning and sequential decision making for multidimensional datapart I: The noiseless case," *IEEE Transactions on Information Theory*, vol. 53, no. 9, pp. 3001–3020, Sept. 2007.

[9] A. Cohen, T. Weissman, and N. Merhav, "Scanning and sequential decision making for multidimensional datapart II: The noisy case," *IEEE Transactions on Information Theory*, vol. 54, no. 12, pp. 5609–5631, Dec. 2008.

[10] V. Krishnamurthy, "How to schedule measurements of a noisy markov chain in decision making?" *IEEE Transactions on Information Theory*, vol. 59, no. 7, pp. 4440–4461, July 2013.

[11] A. Amrico, M. Khouzani, and P. Malacaria, "Conditional entropy and data processing: An axiomatic approach based on core-concavity," *IEEE Transactions on Information Theory*, vol. 66, no. 9, pp. 5537–5547, Sept. 2020.

[12] F. Gorunescu, "Intelligent decision systems in medicine a short survey on medical diagnosis and patient management," in *2015 E-Health and Bioengineering Conference (EHB)*, 2015, pp. 1–9.

[13] J. Yang, J. Zhao, F. Luo, F. Wen, and Z. Y. Dong, "Decision-making for electricity retailers: A brief survey," *IEEE Transactions on Smart Grid*, vol. 9, no. 5, pp. 4140–4153, Sept. 2018.

[14] Y. Rizk, M. Awad, and E. W. Tunstel, "Decision making in multiagent systems: A survey," *IEEE Transactions on Cognitive and Developmental Systems*, vol. 10, no. 3, pp. 514–529, Sept. 2018.

[15] H. Wu, "Multi-objective decision-making for mobile cloud offloading: A survey," *IEEE Access*, vol. 6, pp. 3962–3976, Jan. 2018.

[16] C. H. Weiss, M. J. Bucuvalas, and M. J. Bucuvalas, *Social Science Research and Decision-Making*. Columbia University Press, 1980.

[17] J. Berscheid and F. Roewer-Despres, "Beyond transparency: a proposed framework for accountability in decision-making AI systems," *AI Matters*, vol. 5, no. 2, pp. 13–22, Aug. 2019.

[18] V. Dimitrova, M. O. Mehmood, D. Thakker, B. Sage-Vallier, J. Valdes, and A. G. Cohn, "An ontological approach for pathology assessment and diagnosis of tunnels," *Engineering Applications of Artificial Intelligence*, vol. 90, p. 103450, Apr. 2020.

[19] N. N. Taleb, *The Black Swan: The Impact Of The Highly Improbable*. Random House, 2007, vol. 2.

[20] S. Enayati and H. Pishro-Nik, "A framework for probabilistic decision-making using change-of-probability measures," *IEEE Access*, vol. 8, pp. 159 331–159 350, Sept. 2020.

[21] D. Geman, H. Geman, and N. N. Taleb, "Tail risk constraints and maximum entropy," *Entropy*, vol. 17, no. 6, pp. 3724–3737, June 2015.

[22] J. Rubin, O. Shamir, and T. Naftali, "Trading value and information in MDPs," *Decision Making with Imperfect Decision Makers*, 2012. [Online]. Available: http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.422.9049&rep=rep1&type=pdf

[23] N. Tishby and D. Polani, "Information theory of decisions and actions," *Perception-Action Cycle, Springer*, 2011. [Online]. Available: https://link.springer.com/chapter/10.1007/978-1-4419-1452-1_19

[24] S. Colombo, "Risk-based decision making in complex systems: The ALBA method," in *IEEE International Conference on Industrial Engineering and Engineering Management (IEEM)*, Bali, Indonesia, Dec. 2016, pp. 476–480.

[25] M.-A. Cardin, H. K.-H. Yue, F. Haidong, T. L. Ching, J. Yixin, Z. Sizhe, and H. Boray, "Simulation gaming to study design and management decision-making in flexible engineering systems," in *IEEE International Conference on Systems, Man, and Cybernetics*, Manchester, UK, 2013, pp. 607–614.

[26] Z. Yingchao, "System of systems complexity and decision making," in *7th International Conference on System of Systems Engineering (SoSE)*, Genova, Italy, July 2012.

[27] M. Chennoufi and F. Bendella, "Decision making in complex system," in *5th International Conference on Optimization and Applications (ICOA)*, Kenitra, Morocco, April. 2019, pp. 1–7.

[28] D. Kahneman and A. Tversky, "Prospect theory: an analysis of decision under risk," *Econometrica*, vol. 47, no. 2, pp. 263–292, Mar. 1979.

[29] P. C. Fishburn, "Subjective expected utility: a review of normative theories," *Theory and decision*, vol. 13, no. 2, pp. 139–199, June 1981.

[30] M. Peterson, *An Introduction to Decision Theory: Second Edition*. Cambridge University Press, 2017.

[31] T. M. Cover, "On the st. petersburg paradox," in *IEEE International Symposium on Information Theory Proceedings*, July 2011, pp. 1758–1761.

[32] T. H. Chan, S. Hranilovic, and F. R. Kschischang, "Capacity-achieving probability measure for conditionally gaussian channels with bounded inputs," *IEEE Transactions on Information Theory*, vol. 51, no. 6, pp. 2073–2088, June 2005.

[33] N. Yazdi and K. Todros, "Measure-transformed MVDR beamforming," *IEEE Signal Processing Letters*, vol. 27, pp. 1959–1963, Oct. 2020.

[34] K. Todros and A. O. Hero, "Robust multiple signal classification via probability measure transformation," *IEEE Transactions on Signal Processing*, vol. 63, no. 5, pp. 1156–1170, Mar. 2015.

[35] L. Jin, R. Mesiar, and R. R. Yager, "Melting probability measure with OWA operator to generate fuzzy measure: The crescent method," *IEEE Transactions on Fuzzy Systems*, vol. 27, no. 6, pp. 1309–1316, June 2019.

[36] M. Peterson, "What is a de minimis risk?" *Risk Management*, vol. 4, no. 2, pp. 47–55, Apr. 2002.

[37] H. Swildens and E. Yee. (Feb. 7, 2017) The venture capital risk and return matrix. [Online]. Available: http://www.industryventures.com/the-venture-capital-risk-and-return-matrix/

[38] T. Q. VC. (Aug 23, 2018) How to win in venture capital: focus on the fat tails. [Online]. Available: https://blog.usejournal.com/power-laws-in-venture-capital-why-the-long-tail-matters-22e057c6fa34

[39] Guillem. (May 18, 2017) Understanding the nature of venture vapital returns. [Online]. Available: https://medium.com/@guillemsague/understanding-the-nature-of-venture-capital-returns-95846c65c049

[40] R. Wiltbank. (Oct. 13, 2012) Angel investors do make money, data shows 2.5x returns overall. [Online]. Available: https://techcrunch.com/2012/10/13/angel-investors-make-2-5x-returns-overall/