The Doomsday Machine Revisited: Measuring Systemic Fragility in the Asia Pacific

Sidharth Kaushal

Abstract

In the buildup to the first world war the cabinets of Europe found themselves shackled, in Henry Kissinger’s (1994) memorable formulation to a “doomsday machine” of mobilization timetables which ensured that, almost despite themselves, the great powers of Europe found themselves lurching towards a continent-wide war that would demolish two ancient European houses and leave two others wracked by social and political upheaval. The great puzzle of the first World War is the lack of a theoretical framework within which scholars can accurately describe both the lurch to war and the scope and scale of the war that ensued. While existing theoretical frameworks such as the study of crisis stability, escalation and the offence defence balance all proffer answers to this empirical puzzle, they either answer the question only partially (such as, for example, the likelihood that a crisis might escalate to war as opposed to the scope and scale of the ensuing conflict) or, like the offence defence theory, rest on questionable logical and empirical pillars. The fundamental lacuna is the absence of a simple heuristic by which we might unify both the likelihood of crisis instability escalating to war and the likelihood that such a war might assume general, rather than limited proportions. As such, then, this paper offers a theoretical framework that analyses mutual deterrence through the lens of fragility as opposed to stability. Borrowing from work in the field of finance, the paper argues that scholars of international security should refocus their interest from the likelihood of great power MIDs or the propensity of crises to escalate to the level of warfare, to the study of how well or poorly equipped a system of mutual deterrence is equipped to absorb the shock of military volatility (as opposed to necessarily precluding it, which has been the focus of previous work). Applying this heuristic to the case of the Sino-American strategic balance in the Asia-Pacific, I argue that the primary security threat facing both powers is not potential instability (lamentable though this may be) but rather the fragility of both the states’ military systems. I then conclude with some of the counterintuitive policy implications that may arise from viewing the security architecture of the Asia-Pacific through this lens.
Prediction, as the legendary coach Yogi Berra famously opined, is particularly difficult when it is about the future. The purpose of this paper, therefore, is not to predict the likelihood of politico-military instability in the Asia-Pacific but rather to argue the rather counterintuitive case that it is precisely because of the region’s relative stability that the military balance between its dominant and rising powers is fundamentally fragile and apt to the effects of contagion and escalation. In effect, it is precisely because of our inability to predict rare and improbable events (at least with certitude) that human systems are often rocked and buffeted when such events do occur (Taleb, 2010). The paper will be divided into three parts. The first section will introduce the existing literature on both the stability of strategic balances and the particular conditions prevalent within the Asia-Pacific, along with the limitations of this work. The paper will then articulate the logic and utility of a framework based upon the concept of fragility, as opposed to stability. The second portion of this paper will examine the doctrine of the PLA’s second artillery corps (now the PLA Rocket Force) and the nascent airsea battle concept through the lens of this heuristic and articulate the parallels between the doctrines of the two great Pacific powers and those of the great powers on the eve of the first World War. Finally, the paper will turn to the policy implications that we might derive from this framework.

At the outset, however, it must be clarified that both the interpretation of the two countries’ military doctrines offered here and the broader implications that this paper derives from them are of diagnostic value and must be subject to falsification before any substantive conclusions might be drawn from them. That said, as Max Black proclaimed, every theory “begins with a metaphor and ends with algebra” (Black, 1962. The historical analogies that this paper’s thesis is built upon attempts to approach the “algebra”, but might very well serve as a first cut.

Existing literature about great power conflict tends to focus largely on the capacity of a system to expunge volatility and its effects. The literature might be broadly grouped into two categories - literature that seeks to explain why crises might escalate to the level of warfare and literature that focuses on either the duration of wars or the likelihood that they might escalate vertically from conventional to nuclear conflict. The first category of literature encompasses frameworks such as offence-defence theory, submits that the nature of available weaponry, coupled with factors such as geography, provides incentives to states to either escalate or deescalate conflicts with their rivals to the level of open warfare (Jones, 1991, 660-665). A second vein of literature, based on rationalist bargaining models argues that asymmetries of information and the inability of states to make credible commitments can largely account for the escalation of disagreements to the level of open warfare.

Both strands of theorising suffer from logical and explanatory issues. As has oft been noted, the ability to classify weapons as being inherently offensive or defensive is somewhat circumscribed by
the fact that a good deal of weaponry can abet each purpose, depending on the force structure and doctrine that it is integrated into. Moreover, the value of technological changes that impact factors such as mobility is somewhat indeterminate, inasmuch as one could plausibly contend that mobility favours an attacker (by allowing him to mobilize more quickly than his target) or a defender (who might be able to utilise the mobility of his forces to compensate for the fact that the attacker can choose the time and place of an offensive) (Mearsheimer, 1983,27). Additionally, the utility of the offensive tells us little about the scope of a war. A willingness to mount offensive operations can result in offensive actions of limited scale for limited ends (such as Bismarck’s succession of wars in the build up to German unification or Israel’s campaigns during the wars of 1948 and 1966) or a commitment to the use of force on a vastly larger scale (as was the case in conflicts such as the Napoleonic wars or the first World War).

Rationalist theories such as those of Slantchev(2003) Wagner(2000) and Fearon (1995), by contrast, could plausibly explain the tendency of crises to escalate to the level of warfare and the duration of wars but cannot explain the scope and intensity of conflict. The bargaining model of information transmission could, for example, encompass long lasting, low scale wars of attrition such as the Arab Israeli war of 1967 as easily as it could massive conflagrations such as the World Wars. Moreover, their treatment of relative asymmetries of information as being either exogenous to strategy, or determined by regime type tend to ignore the fact that the capacity of the state to gather adequate information both about the limits of its adversary’s will and capabilities. For example, in the build up to World War I, several diplomatic demarches such as the so called “Willy-Nicky” telegrams between Kaiser Wilhelm and Tsar Nicholas, failed to communicate the two leaders’ mutual desire to either demonstrate resolve or explore avenues for conflict avoidance - in no small measure because the strategies available to both parties required them to commence mobilization before new information could be adequately evaluated (Cimbala,1994,15-30). Moreover, even if escalation to the level of warfare could be explained by informational asymmetries, the sources of these asymmetries is left unclear by existing models. While Slantchev (2003) argues that relative power parity increases uncertainty regarding relative capabilities, it might as easily be argued that relative parity increases certainty on both parties; that war, should it occur, will be highly costly and ex post facto inefficient - a point admitted in Slantchev’s own later work (Slantchev 2004). Given that parties tend to focus not just on whether a war can be won, but also on whether it can be won quickly and cheaply, the logic rationalist arguments appears indeterminate.

A second raft of literature focuses on escalation within war. While a good deal of value has been added to the study of warfare by works on both intended and unintended escalation, the literature has, thus far, remained hamstrung by a focus on the boundary between nuclear and conventional warfare. This ignores the capacity for both horizontal and vertical escalation within the context of conventional wars. Additionally, the literature on escalation tends to assume a process by which the use of force is
gradually ratcheted up as each party seeks to outdo the other. However, this assumption tends to obscure the phenomenon observed during World War I – that of states using “total strategies for limited aims” to employ levels of force that in both cost and risk exceeded the value of the successful resolution of the dispute at stake at the outset of a campaign. While wars such as the first World War in which this is the case are outliers, it is precisely, as will be argued, their rarity and extreme consequences which make them both difficult to account for using linear bargaining models and render them extremely important within the context of the study of international conflict.

To fill this gap in the literature, this paper proposes to repurpose the concept of fragility from the field of finance (among others) to offer scholars and practitioners a simple heuristic by which they may estimate the probability that a rare event will cause a systemwide failure. As outlined by Taleb et al (2012, 3-5) the fragility of a system is defined not as its ability to avoid volatility (such as a liquidity crisis or a military clash) but rather its capacity to retain its essential characteristics and avoid outsized losses in the face of volatility. The more fragile a system is, the more likely it is to demonstrate convexity - that is to say, the harm of a shock is negatively correlated to the rarity of the shock itself. A number of mechanisms underpin this observed tendency of complex systems such as the number of complex interconnections within the system which produce a negative feedback effect (in the event of a market panic, for example) or the tendency of actors to “hide their risks in the tails” i.e to take high risk bets on the positive payoffs to be gleaned from low probability events because the perceived low probability of adverse shifts leads them to take risks that, cumulatively, render the system fragile (Taleb et al,2012,5).

The concept of fragility has its analogue in the field of strategic balancing. Of particular importance is the tendency of actors to adopt risky postures under circumstances of strategic stability precisely because the perceived likelihood of a negative payoff (a catastrophic war, for example) are low while the positive payoffs of an assertive military doctrine are high - both for military organisations and for political leaders who depend on the credibility of their militaries in order to negotiate effectively. When asked about the inflexibility of both the intra European alliance system and rigid mobilization timetables in the wake of a war scare in 1912, Alfred Von Kiderlen-Wachter, the then German Foreign Minister brushed off these concerns, claiming that “war could only happen if one side were so unfathomably foolish or bluff so badly as to be unable to back down” (Coker,2014,8). The oft cited low civilian oversight over military matters (much of Schleiffen’s plan, for example, was only partially known by Germany’s political leadership at the outset of the war) can also be attributed to a lack of oversight of organizations that were reaping the professional rewards of taking major risks (Cimbala,1994-35-40).
Building on Taleb’s work, Johnson and Gheorghe (Johnson, Gheorghe, 2013, 164), have operationalised the concept of fragility yielding three salient criteria – namely:

1) Efficiency vs Risk - In order to increase its efficiency, a system eliminates redundancies and optimises the speed at which it operates. This, however, increases its propensity for a loss of control in the event of an unforeseen event, insofar as it operates too quickly for human operators to control in due time and has eliminated parts that duplicate the same role - rendering it fragile.

2) Interconnectedness - the more interconnected the parts of a system are, the greater the likelihood of contagion, with failures spreading from one part of the system to the next.

3) Stress Starvation - the more starved of volatility a system is, the less likely it is to be prepared for an extreme event

Conceiving of the two parties in a great power strategic balance as being part of a shared system of mutual deterrence, I argue that we can utilise each criterion to assess whether the military balance between two or more great powers is fragile or robust. The efficiency of each power’s military system refers to the speed at which it can mobilise a large proportion of its available force - which in turn, circumscribes the time available to decision makers and increasing their margin for error. Indeed, the tradeoff between efficiency and fragility has already been intuitively grasped by classical military theorists such as Clausewitz, who noted that friction (the myriad countervailing organisational and logistical difficulties that hinder the mobilisation of forces) is a central impediment to war, exceeding its role as a rational instrument and moving towards the ideal of total war. More recently (and, I will argue, presciently) the Soviet Marshal N.Ogarkov claimed that the development of deep strike technologies had removed the element of friction from war - making conventional war comparable, in its destructiveness, to its nuclear equivalent (On War, 190) (Metz, Kievit, 1995, 2).

The interconnectedness of a military system can be measured qualitatively by examining the degree to which the deployment of each of its arms and the deployment of forces in each of the five planes of war (air, land, sea, cyberspace and space) depend upon one another. Additionally, the degree to which the partial mobilisation of forces in a single plane depends upon the mobilisation of other parts of a state’s fielded forces in the same plane is, do some degree, a gauge of its military system’s interconnectedness. For example, in the buildup to the first World War, a partial Russian military mobilisation against Austria but not Germany, was deemed infeasible as the disruption that it would cause to railway timetables would make general mobilisation later, should it be needed, impossible. Additionally, the overlapping duties of military districts meant that mobilisation in a single district naturally entailed mobilisation in neighbouring districts (Cimbala, 1994, 40-60). In particular, if a
single domain such as space or cyberspace is central to the use of forces in all other fields, a system can be deemed highly interconnected and, by extension, fragile.

Finally, the degree of stress starvation is a proximate cause of risk acceptant behaviour - a low level of insurance against improbable negative events which, in turn is a component of system fragility.

In the following section, the paper will articulate the implications of this framework for the Sino-U.S strategic balance in the Asia-Pacific.

The Sino-U.S Strategic Balance

While analysts of the strategic balance in the Asia-Pacific tend to differ over the degree to which one can be sanguine regarding the prospects for the region to accommodate its status quo and rising power, most, if not all, seem to agree with Robert Ross’ assessment that the present strategic balance between the continental and maritime power is a stable one. As David Shambaugh has it, present PLA forces do not evince any desire to effect anything beyond limited deterrence in regional contingencies (Shambaugh,2004). Somewhat more pessimistic authors, by contrast, contend that the PLA’s anti access area denial strategy gives it the capacity to challenge the U.S ability to intervene in local conflicts (such as a contingency involving Taiwan) but do not necessarily see the A2AD strategy as amounting, in and of itself, to a threat to regional stability (Biddle, Oelrich,2016,7-48).

Moreover, to the extent that the stability of the region is called into question by more pessimistic scholars, their focus tends to be on Sino-U.S interactions at the grand strategic level , or on China’s potential to eventually develop a blue water fleet (Friedberg,2011)(Holmes,Yoshihara,2013).

By contrast, this paper will demonstrate that while the Pacific order is stable, it is inherently fragile and liable to the effects of systemic contagion in the event of a major military event.

Counterintuitively, the paper will then argue that the presence of comparably sized Chinese and American surface fleets in the Pacific, coupled with limited volatility in the relations between the two powers, may actually serve the cause of peace in East Asia. In the following sections, the doctrines, force structures and capabilities of both American forces in the Pacific and the PLARF will be appraised along the three criteria of efficiency, interconnectedness and stress starvation.
The PLARF

The Second Artillery Corps Envisions Future War

A central assumption that pervades doctrinal writings both within the PLARF and other branches of the PLA is that future wars will be of a limited nature, in accordance with the long held doctrinal tenet of local war under high tech/informationalized conditions. Moreover, war itself is treated as an improbable event in the context of a largely stable, if competitive, international environment. Indeed, an emphasis is placed upon avoiding direct clashes with an opponent, with one analyst arguing that “since the end of the Second World War, in the majority of wars in which the side with inferior equipment has defeated an enemy with superior equipment, the inferior side has won a relative military victory, compelling the superior enemy to stop fighting or to retreat from the battlefield” (Jiang,1997,115-117).

An emphasis, then, is placed upon strategic signalling, with official publications such as Intimidation Warfare concluding that the judicious use of shore based missiles as warning signals could, without engaging a carrier strike group, force it from the first island chain so as to avoid the risk of a direct engagement (Zhao,2005,188). In a similar vein, the Second Artillery Corps’ most definitive publication, the Science of Second Artillery Campaigns, identifies five non kinetic means by which a CSG can be expelled from China’s littoral:

- Firepower harassment
- Frontal firepower deterrence
- Flank firepower expulsion-encirclement
- Concentrated fire
- Information attacks

Each of these attacks presumes that by raising the risks for carrier strike groups operating in the Asia Pacific, the Second Artillery Corps can exclude them from the region long enough for the PLAN to achieve limited regional objectives. In addition, the authors of the Science of Second Artillery Campaigns discuss the possibility of using EMP warheads to knock out the electronic capabilities of a carrier strike group (Yu,2004,218-395). Similarly, other analysts debate the prospect of cyber attacks on vital C4ISR nodes precluding the involvement of the U.S in a regional crisis (Wang and Zhang, 2000,p. 95). Finally, the Science of Second Artillery Campaigns calls for the creation of well defined no-fly and exclusion zones in the event of a conflict that the Second Artillery Corps could deter CSG’s from entering (Yu,2004,218-395).
Yet there exists a doctrinal lacuna in PLAs thinking about escalation in the Asia-Pacific. Firstly, much of the PLA’s literature emphasises the speed and decisiveness with which forces can be brought to bear in a modern context, coupled with the lethality of the U.S network of C4ISR systems, logistical nodes and bases. Moreover, contrary to the stated objective of war by intimidation, much of the PLA’s conceptual literature highlights the need for rapid escalation in the early phases of a conflict in order to deny the U.S the capacity to operate in the Pacific. As one analyst has it, “in a limited high-tech war, where the pace of action is fast and the duration short, a campaign often takes on a make-or-break character. Clearly the quick and decisive battle assumes much more importance in such a war” (Mulvenon, et al, 2006, 47). Similarly, other analysts are of the opinion that given the time it takes for the U.S to assemble the full panoply of its forces in the Pacific, “this lengthy period of war preparations undoubtedly provides an adversary with quite a few opportunities that it can exploit, by launching a massed attack cutting off supply lines, for instance, causing the enemy to collapse without a battle because it is unable to receive supplies in a timely fashion” (Li, 1995, 190). In effect, avoiding battle implies large scale preemptive strikes on logistical capabilities in the Pacific rather than merely symbolic shows of force. Moreover, given the capacity for limited missile bombardments to be misinterpreted, along with the serious possibility that this would constitute a basis for massive conventional escalation by the U.S, the Second Artillery Corps has a strong incentive to pre-emptively escalate attacks to the U.S logistical and C4ISR capabilities in the Pacific, rather than restricting its operations to symbolic shows of force. Indeed, some of its proposed policies (such as no fly zones) effectively require the suppression of U.S air defences, ground based aircraft and C4ISR capabilities to be effective - a point conceded by the Science of Second Artillery Campaigns which suggests that horizontal escalation may well be acceptable and even de-escalatory in the long run (Yu, 2004, 218-395).

Secondly, PLA authors stress the initial advantages offered to the side that can blind the electronic reconnaissance capabilities of its opponent to deal a decisive conventional blow to his forces, coupled with the capacity for rapid escalation offered by deep strike capabilities. The element of friction, then, is reduced to zero inasmuch as the decisive salvos of a future war are assumed to be those that will be exchanged almost instantaneously after hostilities commence (Zhang, 2006, 616-628). Like Ogarkov, PLA writers perceive a world in which the efficiency of both their own and the U.S strategic forces in the Pacific is rapidly increasing, both due to the ease of mobilising one’s own forces and the speed at which a decisive blow can be struck at the outset of a war due to the integration of modern forces with a clustered network of reconnaissance and communications capabilities.

Clearly, the PLA’s writings on war in the Asia-Pacific demonstrate two features - internal contradictions between a belief in the rapid and decisive use of force and a simultaneous faith in the PRC’s capacity for achieving its aims via limited shows of force or de-escalation after a period of intense military activity. As B. Klein stated, such conceptual schizophrenia typically evinces a rupture
between the dominant strategic narrative of the political elite and the operational realities of military planning (Klein,1988,133-148). This would point strongly towards the likelihood that the operational doctrines of the Second Artillery Corps are highly risk acceptant precisely because the political and military leadership of the P.R.C does not expect a general war to break out in the event of a crisis, but can utilise the deterrent effect of being able to rapidly escalate in the Pacific to good use in a regional crisis short of war.

As such, the conceptual framework within which the Second Artillery Corps operates, reconciles a high concern with efficiency in the early stages of a major conflict, even at the risk of horizontal or vertical escalation, with a belief in brinkmanship short of war. When viewed in the context of the wider belief in the region’s strategic stability, this suggests that the Corps is gambling on the low probability of a general war in which it might have to operationalise its escalatory doctrinal precepts. In effect, this indicates both a lack of redundancy and a lack of stress exposure - both of which contribute to systemic fragility.

**Procurement Policies**

Over the last decade, the Second Artillery Corps has seen its procurement programmes restructured substantially, with a decreased emphasis on nuclear deterrence and an increased emphasis on the purchase of SRBM’s. Of particular interest to Western analysts has been the Corps’ increasing emphasis on the purchase and production of ASCMs such as the Israeli Harpy missile the CJ-10 and the DF21-D (DoD,2013,5-10). However, the C4ISR capabilities to track and target a carrier group at sea have still not been developed in their entirety and those capabilities that have been developed are highly vulnerable to disruption (Cordesman,2016). Indeed, the bulk of the forces’ procurements over the last decade have been DF-21C and DF16 IRBMs that would be utilised to target logistical nodes in the Asia-Pacific as opposed to moving carrier groups U.S China Economic and Security Review Commission-Annual Report to Congress,2016,315). This is not to say that the PLA is not attempting to develop the capacity to target carriers; merely that at the outset of a campaign, it may have little choice but to mobilise the entirety of its forces inasmuch as it lacks the survivable C4ISR capabilities to restrict itself to targeting carrier groups.

Nor do the other arms of the PLA offer much capacity for localised mobilisation. The PLAAF fields over 100 fighter bombers equipped with ASCM’s, but this force is insufficient to deter or swarm multiple carrier groups (Jackson, 2003,1) (OSD, 2004, 35)( Hewson, 2004,1). More importantly, the PLAAF’s more recent purchases over the last decade such as the SU-27 are likely to complement an IRBM attack upon logistical and C4ISR nodes (Cliff et al,2007,50-60)
Nor does the PLAN’s submarine fleet offer a viable alternative as it is largely comprised of noisy diesel submarines that are not likely to fare well against competent ASW capabilities (Jane’s, 2004) (Saunders, 2004).

Finally, we address the PLAN’s growing surface fleet. While the PLAN will likely never match the USN in terms of either the quality or quantity of its vessels, it could, in theory, serve as a “fleet in being” capable of posing a sufficient threat at a local level to reduce the likelihood of U.S intervention in a local conflict involving the P.R.C. This would also allow the PLA a degree of flexibility in the articulation and execution of its war plans—insofar as a sizable PLAN surface force could fulfill the function of a fleet in being sans the preemptive use of force for which the second artillery force is geared. The fleet has seen a number of recent additions including several Sovremenny class destroyers equipped with moskit ASCM’s (Cole, 2010, 100). Additionally, the PLAN is in the process of integrating its existing carrier, the Lioning with its Luyang GMD in order to form a carrier group—a process which it will repeat with several Kuznetsoy class carriers that it is refurbishing (Ellman, Bussert, 2011). That said, the lack of catapult launchers aboard PLA aircraft carriers, coupled with chronic weaknesses in ASW capabilities mean that, for the foreseeable future, the PLAN is unlikely to fulfill the function of a fleet in being (Erickson Denmark Collins, 2012, 30).

Discernibly, the existing procurement policies of the PLA strongly point to a heavy reliance on IRBM’s and SRBM’s that would, in order to be used effectively, have to be mobilised and used relatively early in a conflict and, very possibly, pre-emptively.

**Interconnectedness**

A final aspect of fragility which will be examined in the context of the PLA’s military posture is interconnectedness—namely, the degree to which the mobilisation of military force in one theatre or sphere of war necessitates the mobilisation of force in others.

Given the PLA’s proclivity to argue that information networks constitute the center of gravity of a modern force, it is unsurprising that control of the flow of information in the battlespace and its denial of access to an opponent is a substantial part of the PLA’s thinking about the character of war. What is somewhat less obvious is that fact that, given the current force structure of the PLA, efforts to control an opponent’s access to information by cyberattacks for example, cannot occur in an operational vaccum or be disaggregated from wider war plans as an independent escalatory step (as most analysts tactily or explicitly assume). While in principle the PLA has experimented with the use of ASAT missiles and has, by all accounts, deliberated the use of cheap Kaituozhe satellites as space mines, the ability of the PLA to target the U.S’ C4ISR system in space is limited, not least of all due to the sheer density of the satellite network manned and maintained by the U.S. Additionally,
mobilising ASAT missiles would take a process of weeks - not only giving an opponent ample time to perceive the mobilisation but also leaving him uncertain of whether it was part of a more general mobilisation. This, thereby gives him an incentive to treat it as one and, by extension, makes it impossible for the PLA to selectively mobilise ASATs without mobilising the rest of its missile forces. Conceivably, the most attractive option for a combatant seeking to blind its opponent is to focus on targeting ground stations in conjunction with any attempt to deny an opponent access to the electronic and cyber spaces by non lethal or non kinetic means (Coker,2014,168) (DeBlois et al.,2004,57-60).

Similarly, while cyberattacks may well be used to impede C4ISR systems, the blurred threshold between a cyber and a conventional attack in the eyes of the target state (inasmuch as one may precede the other) coupled with the capacity for states to detect and recover from cyberattacks at relative speed implies that any use of cyberwarfare cannot be disaggregated from the use of conventional forces. Given that an opponent facing a cyberattack may well interpret it as a prelude to a conventional assault and escalate immediately, the PLA has a strong incentive to do so as well in the immediate wake of launching a cyber salvo - both to pre-empt a conventional response and to capitalise on the initial paralysis caused by cyber disruption-(Libicki,Gomert,2014,1-19).

Hence, the need to preclude the U.S from attaining information superiority requires the PLA to engage in massive horizontal escalation that would see it targeting U.S C4ISR systems, logistical nodes and bases on the territory of its allies - thus widening the scope of any war.

Secondly, stated objectives of the PLARF, such as declaring limited exclusion zones requires not only a demonstrated capacity for enforcement but the ability to pre-emptively degrade air and naval assets that might challenge these zones of exclusion (Mueller,2013,7). Moreover, even if the PLARF did intend to utilise ASCMs in and ASBMs exclusively in order to maintain these zones of exclusion, any effort to selectively mobilise ASBMs such as the DF21-D not only risks misinterpretation as a general mobilisation, but is also impossible given the inflexibility of the war plans that we can infer from PLARF exercises. These exercises demonstrate an acute sensitivity to the possibility that the PLA may well find itself operating under circumstances where it has lost control over part or all of its C4ISR and faces the prospect of pre-emption. If indeed this is the case, then it might be assumed that the capacity to direct limited strikes has eroded, whilst the PLA expects to be faced with a “use it or lose it” dilemma with regards to its A2Ad capabilities - giving it strong incentives for horizontal escalation (Kamphausen,Lai,Tanner,2014,322-344. Indeed, it is on these grounds, that PLA analysts have criticised the notion of military gradualism, arguing that in the event of a conflict the capacity for limiting contagion is so circumscribed that the PLA must escalate rapidly and decisively if it is to secure any chance of victory. (Wang,Zhang,2000,96). Of particular importance is the reduction of enemy airfields - with the pre-emptive destruction of an opponent’s deep strike capabilities seen as a
prerequisite to any campaign, be it limited or total in its aims. As such, regardless of the limited aims of the PLA in any given conflict, it must commence the campaign with a general mobilisation of its SRBM and IRBM forces and large scale horizontal escalation against potentially hostile airfields along the first island chain to achieve the prerequisite campaign objective of “disrupting the enemy air strike plans… target the enemy’s command-and-control systems and fuel and ammunition supply systems; if it is aimed at degrading an enemy aviation corps group to reduce the pressures from its air strikes, one should target the aircraft parked on the tarmacs of airports housing the enemy’s main bomber and fighter-bomber aviation corps.”(Wang,Zhang,2000,362).

The individual components of the PLA’s force structure are thus symbiotically linked to one another. In many ways, this phenomenon is not unique to the PLA, insofar as the revolution in military affairs, with its emphasis on the creation of networked military systems that can respond to contingencies rapidly lends itself to the creation of interconnected systems. Given the inherent vulnerability of these systems to a loss of access to C4ISR, coupled with the inherent inflexibility of its posture which, by necessity links the use of force in a domain such as cyberspace to the mobilisation and use of the assets of the PLAAF and second artillery corps, the PLA faces the prospect of internal contagion, whereby the capacity to selectively mobilise parts of the organisation does not exist - rendering it fragile in the face of rare pressure.

Having analysed the fragility inherent in the current PLA system, the paper will now analyse the U.S regional posture through the lens of the same framework.

**U.S Regional Posture**

**Doctrinal Precepts**

At the core of the U.S doctrinal response to the PLA’s development of robust anti access area denial capabilities is the AirSea battle operational concept. The concept revolves around three precepts - withstanding an initial attempt to blind American C4ISR capabilities, seizing the initiative by degrading the PLA’s own C4ISR and deep strikes to degrade both the second artillery forces kill chain (Krepinevich et al,2010,13-15). However, the capacity to withstand an initial and sudden escalation is somewhat limited, as is the capacity of the U.S to redeploy forces to the Pacific theatre adroitly in the face of a crisis in which substantial theatre ballistic missile forces are degrading its access to the theatre and the logistical bases upon which this access depends. Indeed, given the growing capabilities of the second artillery force, which has upgraded its DF21-D class missiles to strike at a range of over 1500km, even access to bases along the second island chain such as Guam may well be contested in the future. It is implicit in the airsea battle doctrine that blinding and degrading the Second Artillery Corps forces pre-emptively are objectives that the U.S must seek at the
very outset of a conflict. Within the context of this operational concept, the U.S must consider pre-
emptive strikes in a preconflict phase if it believes a conflict likely, as it may lack the capacity to
sustain a credible conventional “second strike” in the event that the PLARF initiates hostilities.

While it has been argued by figures such as H.Kazianis (Kazianis,2013,1) that the concept includes
efforts to withstand or defend carrier strike groups and vital logistical and communications, the
untested nature and limited local availability of forces such as theatre missile defences makes it
unlikely that the U.S can afford to passively absorb fire at the outset of hostilities before shifting to
the offensive.

Indeed, the AirSea battle office’s official articulation of its posture records as much: “indications or
warning, forward friendly forces will be in the A2/AD environment at the commencement of
hostilities. As a result, the steady state posture and capabilities of forces must be able to provide an
immediate and effective response to adversary A2/AD attacks through high tempo operations in the
A2/AD environment. Additional forces introduced into the threat environment should be able to
promptly integrate into the existing force posture”. The paper goes on to reiterate that attacks in depth
must remain the primary operational means by which the U.S protects its fielded forces and enables
fresh forces to be relayed into the theatre of operations (AirseaBattle Office,2013,4-5).

To this end, the U.S has redeployed several carrier strike groups to the Asia Pacific and is in the
process of overhauling the USAF assets positioned in airbases along the first island chain, with
existing F-35 and F-22s making way for longer range B-2 bombers and research and development
being refocused on the development of more modern deep penetration bombers
(Biddle,Oelrich,20161-15). Moreover, the AirSea Battle concept has stimulated the development of
cyberwarfare capabilities that to all effects and purposes seek to emulate the PLA’s efforts to seek
information superiority at the outset of a conflict.

As has been noted by several commentators, the ramifications of the Airsea battle concept make it
difficult to operationalise in all but total circumstances, creating no small degree of dissonance
between grand strategic ends and strategic means (Hammes,2012,1-5). Perhaps more importantly,
however, is the fact that it leaves a short window of opportunity at the outset of a conflict for the
document to be operationalised - raising the possibility that human error, among other risks, might
dictate the course of a military dispute. The more efficient a system and the less friction (in
Clausewitzian terms) it faces, the smaller the temporal gap between decision making errors and
adverse outcomes.

The first salient point to note is the importance attached to eliminating fog and friction and increasing
relative efficiency. Indeed, early American theorists of the RMA in many ways anticipated the PLA
in the belief that fog and friction could be entirely eliminated from warfare by the efficiency of networked C4ISR capabilities and precision guided munitions (Benbow,2004,90-91). Indeed, it was by observing American conduct in the Gulf War that the PLA reached its conclusions regarding the speed and intensity of modern warfare (Cliff et al,2007,20-50. To all intents and purposes, both forces effectively mirror each other in their willingness to trade the risk of human failure for efficiency - leading to a substantially higher level of systemic fragility in the event of an unforeseen crisis, given the capacity of perfectly efficient systems to escape political control.

The development of Airsea battle, which has been described as occurring in a strategic void, based purely on the logic of military annihilation and which is incompatible with both the need to limit conflict and restore a modus vivendi in after the termination of one (Hammes,2012,1-5). In large measure, the adoption of the concept, is therefore predicated on the belief that conflict will not occur (and concomitantly, the belief in the value of an escalatory warfighting doctrine in manoeuvring and deterrence short of conflict) coupled with a lack of political oversight over military planning in the face of relative stability - which typically allows militaries to adopt their own favoured (and typically offensive) warfighting plans (Posen,1984-10-25). As such, the relative stability of the Pacific has led policymakers to “hide their risks in the tails”- betting on the improbability of a highly unlikely event to gain strategic leverage much in the way the PLARF is betting on the improbability of a crisis severe enough to cause a war when it attempts to gain the coercive leverage that having a credible A2AD capability provides. In effect, this amounts to an absence of volatility producing a highly fragile system.

Finally, it is worth noting the integration of American forces in the region. Aside from the fact that, like the PLA, the U.S necessarily links nonkinetic escalation such as cyberwarfare to operationalising warfighting plans for high intensity conventional conflict. More importantly, given the integration of local country commands into PACOM and in light of the relative lack of dispersion of American forces in forward deployed bases, the mobilisation of forces in one area of the pacific theatre is , for both organisational and strategic reasons, linked to general mobilisation across the theatre. Organisationally, the integration of country commands into a wider pacific command lends itself to theatre wide mobilisation plans while the interrelated nature of the U.S bases, coupled with the risks posed to forces based in a specific allied country in the event that U.S forces elsewhere in the theatre are mobilised makes flexible mobilisation and targeting plans infeasible (Burgess,2016,29)

The existence of this posture satisfies all the criteria for a fragile military system. An added dimension of fragility is added to the conflict by the alliance politics of the region and the fact that the mobilisation of theatre level U.S forces depends on bases and logistical nodes that run across several countries - creating substantial capacity for political contagion(Friedberg,2014). However, given the perceived infeasibility of available alternatives such as blockades ( which are perceived as taking too
long to have political a political effect), the ASB has remained a central component of the wider JAM-GC concept (Kazianis, 2015).

**Systemic fragility and World War I**

The central point that the preceding section has articulated is that an excessive emphasis on regional stability, coupled with efforts to assess the conventional balance of a region by quantifying those platforms deemed offensive, tends to obscure the fact that regional order depends, in no small measure, upon the capacity of a system to absorb volatility rather than preclude it. Thus, quantifying the number of “defensive” shore bases, missiles or offensive ships possessed by one actor, or to estimate the likelihood that they will calculate that they can fight or win a war allows only a partial insight into regional dynamics. Of far greater consequence is the question of whether, in the event of a rare crisis to which leaders commit military force (however improbable such a crisis may be perceived to be) the system of mutual deterrence constructed by both actors is robust enough to preclude a loss of political control over the instrument of war and to prevent war lurching towards its absolute.

Gauging the dynamics of a fragile balance with reference to a historical analogy is a task that one should approach with caution. To paraphrase Tolstoy, every unhappy situation is unhappy in its own way. Nonetheless, given the relative rarity of “black swan” events that reveal systemic fragility, it is instructive to explore the parallels between the Sino-U.S strategic balance and the unravelling of the nominally balanced and defence-dominated European security system during the first World War.

The buildup to World War I saw the emergence of an unquestioned “cult of the offensive” among military professionals in an uncritical belief in the efficacy of offensive military doctrines and, perhaps more importantly, the idea that wars could be won swiftly by mobilising and manoeuvring more rapidly than one’s opponent - in direct contravention of the lessons of the recent conflicts such as the American war and the Russo-Japanese wars (Van Evera, 1984).

The survival of this article of faith was, in no small measure, helped by low levels of political oversight which allowed general staffs to adopt highly risky warfighting plans even as political leaders sought to benefit from the deterrent value provided by these plans along with the fielded forces that underwrote them, whilst remaining confident that war could be either avoided or localised. Indeed, it is worth noting that the German chancellor Bethmann Holwegge maintained this particular article of faith up until the very outbreak of the first World War (Jarausch, 2012) (Zeigler, 1977, 25). Nor was he alone - London’s bond markets did not see prices drop even on the very eve of war (Ferguson, 2006).
Secondly, the bureaucratization of European armies tethered the cabinets of Europe to an mobilisation timetables that frayed the capacity for political control over military events. Given the pace of mobilisation, coupled with the fear that even small delays in mobilisation could have large ramifications, states were effectively bound to escalate crises regardless of the political will to control them. Moreover, the capacity for human error, biased reasoning by a given leader and other human errors would, within the context of such a system, be difficult if not impossible to correct. In effect, the cabinets of Europe were victims of their own military efficiency, which eliminated redundant steps in the build up to war and in the process, made the event of a crisis cascading beyond the level of limited war all but inevitable. For example, even as Tsar Nicholas and Kaiser Wilhelm exchanged the famous “Willy-Nicky“ telegrams, each leader was pushed by his military to order general mobilization in fear of a type II error – namely that a failure to mobilise fast enough would allow the other side to steal a march on it (Albertini, 2005, 559).

Finally, the integration of each military system into both a set of rigid war plans that made the operation of individual parts independently of the system as a gestalt impossible, coupled with the integration of each great power into a system of alliances, made the risk of contagion in the form of both horizontal and vertical escalation all too likely. When Tsar Nicholas attempted to explore the possibility of partial mobilisation against Austria-Hungary but not Germany, in the wake of the Archduke Franz Ferdinand’s assassination, he was informed by General Sukhmolinov that this would be an impossibility given that railway timetables were geared towards general mobilisation and that any change of plans would disrupt them. Moreover, given Russia’s system of interlocking military districts, mobilisation in one district necessarily entailed mobilisation in others to provide logistical support and to bring the army to a war footing. Germany, for its part was also incapable of mobilising selectively or avoiding an invasion of Belgium (which would widen the scope of the war by involving Britain) insofar as the Schleiffen plan depended on siezing and maintaining the initiative and the success of its individual components depended on them all being activated sequentially in short order.

Simply put, then, it was fragility and not instability that led Europe to the first World War. Indeed, as Richard Ned Lebow convincingly argues, European leaders were broadly correct about the general stability of the international system and the low probability that it might result in a war (Ned Lebow, 2014). However, it was precisely this stability that led to the construction of a fragile strategic balance, as leaders discounted the tail risks of an improbable event.
Policy Implications and Conclusion

Paradoxically, had Europe seen more limited wars in the build up to the first World War, it might well have averted a continent wide disaster. While not advocating for limited conflict of course, this paper does advocate the counterintuitive position that limited strategic volatility may well be beneficial to the Asia-Pacific. Specifically, I examine two types of volatility—the buildup of the PLA’s surface fleet and limited standoffs both in disputed maritime waters. While figures such as Holmes and Yoshihara(2013) argue that a symmetrical PLAN buildup (particularly when coupled with what they interpret as the PLAN’s Mahanian strategic ethos) represents a threat to regional stability, it is noteworthy that the naval forces of the two powers are perhaps the only arms of their respective militaries that could, in theory, keep war restricted to a specific domain (the maritime theatre in question) even as both parties were mutually deterred from escalating conflict against each other’s logistical and supply nodes for fear of retaliation or unforeseen escalation. Of course, as was noted above, the PLAN surface and submarine fleet are at present a far cry from being able to challenge the U.S symmetrically - making any tacit agreement to localise war to the maritime sphere inconceivable in the event that war should break out. However, as the PLAN does develop as an arm capable of acting independently of its A2AD shield and acting as, at the very least, a fleet in being or a risk fleet vis a vis the stronger U.S Navy, it becomes conceivable that both powers might mutually agree to disaggregate naval warfare from information operations and the use of land based missile and fixed wing aircraft.

Secondly, I argue that volatility such as standoffs in the South China Seas, coupled with each nation’s use of limited cyber espionage against the other, serves a useful purpose in raising policymakers’ estimation of the probability of a rare outcome and giving them incentives not to gamble against a black swan event.

Finally, developments such as proposed “normalisation” of the status of Japan, coupled with arms sales by the U.S to Taiwan and even ill conceived arguments regarding the nuclearization of South Korea arguably destabilise the Asia Pacific as a system, but also decouple, to a certain degree, regional actors from the “hub and spoke” model - raising the risk of regional conflagrations but lowering the fragility of the system by making it at least conceivable that these conflagrations might, at some point in the future, be localised and carried out absent the need for the sort of military commitment by the U.S that would force it to rapidly escalate to levels of force that would be incommensurable with the political object at stake.

This paper has, therefore, attempted to demonstrate the value of a low cost heuristic with which we might measure the fragility of a system to the assessment of strategic balances in general and the Asia-Pacific in particular. While almost entirely diagnostic in its nature, the paper has outlined
definite criteria by which systemic fragility can be measured, demonstrated (at least in a preliminary sense) the way in which these criteria are applicable to military systems and the counterintuitive policy prescriptions that they might lead us to. In the final instance, however, this paper represents a “first cut” at repurposing the concept of fragility to study outlier events in the realm of security and the concepts outlined here will require substantially more empirical research before they can be regarded as conclusive.


Cole, B. 2010. The Great Wall at Sea. Naval Institute Press:


