A Novel Approach for Mitigating the Effects of Pre-rolling/Pre-sliding Friction on the Settling Time of Rolling Bearing Nanopositioning Stages using High Frequency Vibration

Xin Dong, Deokkyun Yoon, Chinedum E. Okwudire*

Department of Mechanical Engineering, University of Michigan, 2350 Hayward Street,

Ann Arbor, MI 48109, USA

*Corresponding author. Tel.: +1 734 647 1531; Fax.: +1 734 647 9379

E-mail address: okwudire@umich.edu (C. E. Okwudire)

Abstract

Nanopositioning stages are used for ultra-precise positioning in many different applications. Among the bearing options available for long-range nanopositioning stages, rolling bearings are the most cost effective. However, rolling bearing nanopositioning (RB-NP) stages are well known to take a long time to reach their target positions in point-to-point positioning tasks due to pre-rolling/pre-sliding friction. This paper presents vibration assisted nanopositioning (VAN) – a novel approach for reducing the settling time of RB-NP stages using high frequency vibration (aka dither). VAN synergistically combines the electromechanical design and control of RB-NP stages to mitigate pre-rolling/pre-sliding friction using dither, without jeopardizing positioning precision. It is shown in simulations to exhibit superior positioning performance and robustness compared to model-based friction compensation methods. Experiments carried out on a prototype stage incorporating the VAN approach show a 52% reduction of mean settling time, based on 50 trials, when

compared to a case without VAN. Moreover, the heat and potential wear induced by dither are shown to be practically insignificant.

Keywords: pre-rolling/pre-sliding friction, dither, nanopositioning, rolling bearing, repetitive control

1. Introduction

Nanopositioning stages are used for ultra-precise positioning in a wide range of applications, such as MEMS inspection, wafer scanning, micro-/nano-machining and scanning probe microscopy [1], [2]. They can be constructed using flexure, magnetic, fluidic or rolling bearings. Flexure bearings are compact, low-cost, and wear and friction free. However, they are not suitable for long-range, medium-to-high-payload positioning applications due to their limited stroke and load capacity [3], [4]. Magnetic bearings (i.e., maglev) are capable of long-range nanopositioning with excellent off-axis stiffness and translational speeds [5], but their commercial use is currently relegated to the highest end stages (such as those used in semiconductor manufacturing [1]), because of their high costs and complexities. Fluidic (i.e., hydrostatic [6] or aerostatic [2]) bearings are also capable of long-range nanopositioning, but they are not suitable in clean room and/or high vacuum environments.

Rolling bearings, such as crossed roller and linear ball bearings, are the most cost effective bearing type used for long-range nanopositioning. Moreover, they are currently the only commercially viable option for a growing number of long-range nanopositioning applications which must be performed in high or ultra-high vacuum environments (e.g., electron microscopy, thin-film sputtering and focused ion beam [7]). However, rolling

bearing nanopositioning (RB-NP) stages experience nonlinear rolling/sliding friction arising from their inherent rolling elements, end seals/wipers and lubricants [8]. This nonlinear friction adversely affects the precision and speed of RB-NP stages in both the macro-displacement (i.e., gross rolling/sliding) and micro-displacement (i.e., pre-rolling/pre-sliding) regimes of friction [9] – [12]. In the gross rolling/sliding regime, friction is mainly a function of velocity, and stick-slip is a common phenomenon that occurs due to the highly nonlinear behavior of sliding friction at velocity reversals [11]. On the other hand, in the pre-rolling/pre-sliding regime, friction behaves as a nonlinear spring due to elastoplastic deformations and micro-slip of the rolling elements and seals/wipers [8] – [11]. Pre-rolling/pre-sliding friction becomes dominant as RB-NP stages get within microns of their desired positions, resulting in much (e.g., 5 – 10 fold) longer settling times than equivalent frictionless stages in point-to-point positioning [12] – [14]. Such long settling times severely hamper the throughput of the processes for which RB-NP stages are used.

The state of the art in addressing friction in RB-NP stages is to perform feed forward or feedback friction compensation using controllers. In feed forward friction compensation, the friction force experienced by a stage is predicted and canceled out preemptively using a model of friction combined with knowledge of the motion commands applied to the stage [15]. Tracking performance can be greatly improved using feed forward friction compensation if the friction models employed are sufficiently accurate [16] – [18]. However, the methods suffer from robustness problems due to difficulties in accurately modeling friction and the variability of on-machine frictional dynamics [15]. Moreover, the effectiveness of feed forward friction compensation methods on improving settling

performance in point-to-point positioning is hampered by the fact that the reference velocity is zero during settling [13], [19]. Feedback friction compensation approaches make use of actual states (e.g., position and velocity) of the system to improve disturbance rejection, using high-gain feedback, gain scheduling controllers, friction observers, etc. [9], [12] – [15], [20]. However, the use of high-gain feedback makes systems to be prone to instability and/or undesirable limit cycles due to sensor noise, system parameter variation, higher order dynamics, etc. [15]. Feedback methods that employ a model of friction (e.g., observer-based methods, gain scheduling controllers, etc.) suffer from robustness issues due to the variability of friction. Several methods have been proposed in the literature for improving the robustness of model-based friction compensation methods through model parameter adaptation [19], [21]. However, the convergence of adaptation schemes is often unreliable and slow because the identification signals are not rich (or persistent enough) [15].

High frequency vibration (or dither) is well-known to be an effective, robust and model-free approach for mitigating nonlinear phenomena like friction, backlash and hysteresis [22] – [24]. With regard to friction mitigation, dither is often used to reduce stick-slip and friction induced oscillations, e.g., in spool valves of hydraulic servos, disc brakes and flexible robot arms [15], [25], [26]. It has also been used in micron-level-precision positioners, for example, to reduce hysteresis in piezo actuators [27] and to mitigate stiction and Coulomb friction in sliding bearings of machine tools [28]. However, to the authors’ best knowledge, dither has not been used for mitigating pre-rolling/pre-sliding friction in nanopositioning stages, due to its tendency to cause unwanted vibration which jeopardizes nanometer-level precision [28], [29]. Heat and wear are also concerns

that must be dealt with in using dither for friction mitigation in precision or ultra-precision positioning applications [28], [29].

This paper (and its preliminary version [30]) presents vibration assisted nanopositioning (VAN): a novel, effective, robust and model-free method for mitigating pre-rolling/pre-sliding friction, and thus improving settling time, in RB-NP stages using dither. The novelty of VAN and the key contributions of this paper are in: (i) proposing a mechatronic concept where the electromechanical design of a RB-NP stage is synergistically combined with active vibration control to allow dither to be applied without, in theory, vibrating the stage; and (ii) designing a prototype RB-NP stage based on the VAN concept, and experimentally demonstrating significant reductions in settling time using dither without jeopardizing nanometer-level precision, or causing excessive heat and potential for wear.

This paper is organized as follows: In Section 2, the slow settling problem of RB-NP stages is introduced, and the performance, robustness and stability challenges of traditional controllers and model-based friction compensating controllers in addressing it are discussed and demonstrated using simulations. The concept of VAN is then presented in Section 3, and its superior performance, robustness and stability are demonstrated via simulations; a control scheme for handling parasitic vibration that results from non-idealities in VAN is also presented. The design of a prototype VAN stage and experimental validation of VAN in reducing the settling time of RB-NP stages without causing excessive heat and potential for wear is then presented in Section 4, followed by conclusions and future work.

2. Problem Description

2.1. Slow Settling Problem of RB-NP Stages

Consider a servo-controlled RB-NP stage of mass m depicted in Figure 1. Given a point-to-point positioning command, x_r , the actual position, x_s , of the stage must reach x_r as quickly as possible while overcoming friction force F_f , which is dominated by pre-rolling/pre-sliding friction as x_s gets very close to x_r [12] – [14].

Figure 1. Model of a servo-controlled RB-NP stage under the influence of friction.

In its pre-rolling/pre-sliding regime, the force of friction is well known to be displacement dependent, and can be described by a rate-independent hysteresis function [8] – [11], [19]. There are several models with various levels of complexity for describing pre-rolling/pre-sliding dynamics (e.g., Dahl [31], LuGre [20], Generalized Maxwell Slip [32], etc.). However, the Dahl friction model is adopted in this paper because it is simple yet it captures the primary cause of slow settling behavior in RB-NP stages, namely, the nonlinear spring-like characteristic of pre-rolling/pre-sliding friction [12], [13], [18]. The Dahl model is given by the differential equation

$$\frac{dF_f}{dx_s} = \sigma_0 \left(1 - \frac{F_f}{F_c} \operatorname{sgn} \dot{x}_s \right)^i \quad (1)$$

where F_f is the force of friction, σ_0 is the initial contact stiffness, F_c is a measure of the Coulomb friction force and i is a shape factor which is usually set to 1. According to the Dahl model (and other models), there are two key problems associated with pre-rolling/pre-sliding friction which cause the slow settling problem in point-to-point positioning, and create challenges for addressing it using purely control-based techniques:

Problem I: When the displacement magnitude, $|x_s|$, is very small ($< \sim 100$ nm), friction behaves like a quasilinear spring with a very large stiffness coefficient, σ_0 [8]–[11]. This stiff spring makes it difficult and slow for a controller with conservative gains to push the stage into its target position [9], [10], [12].

Problem II: As $|x_s|$ increases, the spring stiffness of pre-rolling/pre-sliding friction decreases nonlinearly and drastically from σ_0 to a value of zero as F_f approaches F_c , implying a nonlinear time-varying system, which is hard to control.

2.2. Challenges of Traditional Controllers in addressing Slow Settling Problem

The most basic and industry standard approach for controlling motion stages is to employ a PID-type controller (e.g., PID, P-PI, PI-P, etc.) with velocity and acceleration feedforward [9], [33]; it is preferable to keep the gains of the controller fixed, to maintain a linear time-invariant (LTI) controller. To deal with Problem I, such a controller must have high integral gain to enable it to quickly overcome the large stiffness of pre-rolling/pre-sliding friction and drive the stage to its target position [9], [34], [35]. However, it is well-known that high-gain feedback controllers are prone to overshoots, limit cycles and

instability in the presence of higher order dynamics, measurement noise, variations in system parameters, etc. [15]. Problem II exacerbates the issues introduced by using high integral gain to address Problem I, because of the rapid and nonlinear changes in stiffness (i.e., system parameter variation) [35], [36].

To further illustrate these problems and provide a reference for subsequent comparisons, simulations are conducted using the RB-NP stage model of Figure 1 controlled using a PID-type controller with proportional (K_P), integral (K_I) and derivative (K_D) gains. The parameters used for the simulations are those of Aerotech ALS-130H RB-NP stage, reported in a recent paper by Bucci et al. [14] and summarized in Table 1. Two step commands with magnitudes 50 nm and 500 nm are used to simulate the settling performance of the stage in the quasilinear and nonlinear regions of pre-rolling/pre-sliding friction, respectively. In each case a $\pm 20\%$ settling window is used, to facilitate easy comparison. The benchmark settling behavior of the stage to the desired positions is taken as the case where $F_f = 0$ and $K_I = 0$ (because integral action is not needed for settling in the absence of friction). Notice that when F_f (given by the parameters of the Dahl model in Table 1) is applied, the stage cannot settle into the desired windows with $K_I = 0$ in both cases due to the very large initial contact stiffness ($\sigma_0 = 8 \text{ N}/\mu\text{m}$) of friction. The overall settling performance can be improved by increasing K_I , as evidenced by the fact that the best settling time (in response to the 50 nm step command) for K_I values ranging from 0 to $10^4 \text{ N}/(\mu\text{m}\cdot\text{s})$, which occurs at $3 \times 10^3 \text{ N}/(\mu\text{m}\cdot\text{s})$, is 33% shorter than that of the benchmark case. However, using such a high value of K_I to reduce settling times adversely affects the stability margins of the control system (as indicated by the oscillations of the displacement signal when $K_I = 3 \times 10^3 \text{ N}/(\mu\text{m}\cdot\text{s})$ in Figure 2 (a)). This loss of stability margins could

easily lead to performance degradation or instability if friction behavior varies or a larger magnitude reference signal is used, as can be seen from Figure 2 (b). In response to the 500 nm step command, the stage suffers from a very large overshoot when using $K_I = 3 \times 10^3$ N/($\mu\text{m}\cdot\text{s}$) (which worked very well for the 50 nm step), resulting 31% longer settling time than the benchmark case. For both step commands, increasing K_I to 6×10^3 N/($\mu\text{m}\cdot\text{s}$) causes stability issues, as indicated by the limit cycle in Figure 2 (a) and unstable response in Figure 2 (b). Therefore, in practice, when using an LTI controller, K_I is tuned conservatively, leading to very long settling times relative to frictionless stages [12] – [14].

Figure 2. Settling results of a PID-type LTI controller in response to (a) 50 nm, and (b) 500 nm step commands using different integral gains. In benchmark cases, no friction is considered and $K_I = 0$. (Unit of K_I is N/($\mu\text{m}\cdot\text{s}$)).

Table 1. Parameters of Aerotech ALS-130H RB-NP stage used in simulations, obtained from Bucci et al. [14].

Parameter	Value
m	1.5 kg
K_P	0.8 N/ μm
K_D	8.5×10^{-4} N·s/ μm
σ_0	8 N/ μm
F_c	1 N

2.3. *Model-based Friction Compensating Controllers and their Challenges*

A lot of work has been reported in the literature on advanced feedback friction compensation methods for addressing the slow-settling problem discussed and demonstrated in Sections 2.1 and 2.2, e.g., high-gain feedback, gain scheduling controllers, friction observers, etc. [9], [12] – [15], [34] – [36]. Due to the sheer physics of Problem I, in order to reduce settling time, these advanced friction compensating controllers also need high integral gains to aggressively overcome the high spring stiffness in the quasilinear regime of friction [9], [13], [14], [34] – [36]; so they share the problems of traditional LTI controllers in this respect. However, they address Problem II by using implicit or explicit knowledge (or models) of friction dynamics to modify their control structure or gains as friction progresses from pre-rolling/pre-sliding to gross rolling/sliding [9], [13], [14], [34] –

[36]. As a result, they avoid some of the pitfalls of maintaining high integral gain during abrupt nonlinear transitions of friction.

However, switching or modulating of controller gains (continuously or discontinuously) as a function of changing friction dynamics brings with it some problems. Firstly, switching itself can introduce stability issues (even if the systems between which switching occurs are by themselves stable) [37]. Stability issues arising from switching are mitigated in some friction compensating controllers by using conservative gains or by introducing dwell time between switching instances, at the expense of positioning time (e.g., [9], [36]). Secondly, friction is extremely variable and difficult to model accurately (especially at the micro scale). Variation of friction can create performance or stability issues for such switching controllers. Adaptive control can enable a model-based friction compensating controller to handle varying friction dynamics [38]. However, as mentioned in Section 1, adaptive controllers suffer from convergence issues related to poor persistence of excitation from smooth reference commands commonly used in precision positioning [15], [38].

To exemplify some of the robustness and stability issues of advanced friction compensating controllers due to their implicit or explicit dependence on a model of friction, let us consider the Nonlinear Integral Action Settling Algorithm (or NIASA for short), recently proposed by Bucci et al. [13], [14]. NIASA is a controller that replaces the fixed integral gain, K_I , of a regular PID controller by a nonlinear function, $K_I(x_s)$, given by

$$K_I(x_s) = \frac{1}{\tau_d} \left[\sigma_0 \left(1 - \frac{F_f}{F_c} \operatorname{sgn} \dot{x}_s \right) + K_P \right] \quad (2)$$

where τ_d is a design time constant. Notice that $K_f(x_s)$ includes a nonlinear portion that varies in proportion to the effective spring stiffness of the Dahl friction model of Eq. (1). Thus, NIASA is representative of friction compensating controllers in which the integral gain is continuously switched, following the Dahl friction model.

Figure 3 shows the simulated settling responses of NIASA to the same 50 nm and 500 nm step commands considered in Section 2.2, using all the parameters given in Table 1 with $\tau_d = 4$ ms [14]. A multiplicative uncertainty parameter $\delta \in [-1, \infty)$ is introduced by Bucci et al. [14] into the contact stiffness, σ_0 , of the Dahl model described by Eq. (1) to help demonstrate the robustness of NIASA to changing friction; i.e., $\sigma_0 = \sigma_{0,n}(1 + \delta)$, where $\sigma_{0,n} = 8$ N/ μ m is the nominal value of σ_0 . It is observed that NIASA is indeed stable within the uncertainty parameter band $\delta \in [-0.5, 1]$ that is presented in [14]. However, notice from Figure 3 that its settling performance varies significantly as δ deviates from its nominal value of zero. Notice also that the system with NIASA experiences severe oscillations for $\delta = -0.5$, and suffers from a limit cycle with large amplitude for $\delta = -0.65$, causing it to not settle within the desired ± 10 nm window, as shown in Figure 3 (a). Comparing Figure 3 (a) and (b), one observes some benefits of modulating NIASA's integral gain in terms of improving stability and performance in the nonlinear regime of pre-rolling/pre-sliding friction. Notice however that the responses of the system in Figure 3 (b) (e.g., its overshoot and settling time) are still highly affected by changes in δ .

These problems of high integral gain, robustness and the need to have an accurate knowledge/model of friction behavior are major challenges that limit the practicality of model-based friction compensating controllers in general.

Figure 3. Settling results of NIASA [14] in response to (a) 50 nm, and (b) 500 nm step commands using different values of uncertainty δ in the initial contact stiffness of the Dahl model.

3. Proposed Solution: Vibration Assisted Nanopositioning

3.1. Concept of Vibration Assisted Nanopositioning

High frequency vibration (or dither) is well known to be an effective, robust and model-free approach for smoothing nonlinearities like friction, hysteresis, backlash, etc. [22] – [28]. Under sufficiently high frequency and high amplitude dither, undesirable nonlinear frictional phenomena like stick-slip and stiction become linearized in an averaged sense [23], [24], [28], as is demonstrated in the Appendix A for pre-rolling/pre-sliding friction in RB-NP stages. A major challenge that prevents dither from being used for friction

mitigation in RB-NP stages is the concern that the applied high frequency and high amplitude vibration could jeopardize the precision of the stage [28], [29]. Moreover, when high amplitude or high frequency dither is maintained for prolonged periods of time, it could cause accelerated wear of mechanical components and wastage of power (resulting in excessive heat generation [28], [29]).

This paper proposes a novel approach for applying dither to RB-NP stages that realizes the friction mitigation benefits of dither while significantly reducing its deleterious effects on the precision of the stage. Figure 4 illustrates the proposed method, which is called vibration assisted nanopositioning (VAN). It has the following features:

- i. Each rolling bearing is not directly/rigidly attached to the stage; rather it is attached to the stage using a compliant joint (e.g., a flexure joint). The joint provides sufficient compliance in the motion direction of the stage while remaining stiff in other orthogonal directions (i.e., the off-axis stiffness of the compliant joint is comparable to that of the rolling bearing).
- ii. The dithering force (F_d) is applied directly to each bearing using a small actuator (e.g., a voice coil or piezo stack) to create sufficient vibration amplitudes at the location of friction (F_f).
- iii. **In ideal situations,** F_d is applied with a phase difference of 180° to the bearings on opposite ends of the stage such that their reaction forces transmitted to the stage cancel out.
- iv. In non-ideal situations, parasitic vibration resulting from any un-cancelled reaction forces is minimized by regulating F_d and/or by applying a compensating force (F_{comp}) to the stage using a control system.

VAN is therefore a mechatronic solution that synergistically combines electromechanical design and active vibration control to ensure that the benefits of dither are realized with little or no parasitic vibration of the stage. Applying dither directly to the bearings through a compliant joint rather than to the entire stage reduces the amplitude and frequency of F_d , and the power required for dithering, which helps reduce heat and wear. This is partly because, when dither is applied directly to the stage, higher amplitudes are required at the stage level to achieve sufficient motion at the bearings, due to low-pass filtering effects of the stage dynamics [15]. Moreover, when dither is applied to the stage without vibration compensation, ultrasonic dithering frequencies (~ 30 kHz) are required to help reduce stage vibration amplitudes to micron levels, without affecting the effectiveness of dither [28]; note that much higher dither frequencies would be needed to achieve nanometer-level vibration amplitudes. However, due to the vibration compensation strategy built into VAN, extremely high frequencies are not required to avoid jeopardizing nanometer-level positioning precision.

Note that, even though the VAN concept is depicted using a stage with one rail and two bearings in Figure 4, it is applicable to other stage configurations (e.g., a stage with two rails and four bearings).

Figure 4. Schematic of the proposed VAN approach for applying dither to a RB-NP stage.

3.2. Simulation-based Evaluation of VAN

3.2.1. Performance and Robustness of VAN (Ideal Case)

Figure 5 shows a simple three-mass model of a VAN stage (closed-loop controlled through the actuation force F_a). Each compliant joint connecting the stage of mass m_s to each bearing of mass m_{bi} ($i = 1, 2$) is modeled by a spring with stiffness k_i and a damper with viscous damping coefficient c_i . Dither is applied in the form of harmonic excitation forces F_{di} , with amplitudes A_i and frequency f . Without loss of generality, the compensating force, F_{comp} , is assumed to be applied at bearing 1; however, F_{comp} could equally be applied at bearing 2, or elsewhere on the stage (e.g., through the stage's actuation force, F_a).

The system dynamics is described by

$$\mathbf{M}\ddot{\mathbf{u}} + \mathbf{C}\dot{\mathbf{u}} + \mathbf{K}\mathbf{u} = \mathbf{F} \quad (3)$$

where $\mathbf{M}$, $\mathbf{C}$ and $\mathbf{K}$ are respectively the mass, damping and stiffness matrices, while $\mathbf{u}$ and $\mathbf{F}$ are the displacement and force vectors of the system. They are given by

$$\begin{aligned}
\mathbf{M} &= \begin{bmatrix} m_{b1} & 0 & 0 \\ 0 & m_{b2} & 0 \\ 0 & 0 & m_s \end{bmatrix}, \mathbf{C} = \begin{bmatrix} c_1 & 0 & -c_1 \\ 0 & c_2 & -c_2 \\ -c_1 & -c_2 & c_1 + c_2 \end{bmatrix}, \\
\mathbf{K} &= \begin{bmatrix} k_1 & 0 & -k_1 \\ 0 & k_2 & -k_2 \\ -k_1 & -k_2 & k_1 + k_2 \end{bmatrix}, \mathbf{F} = \begin{bmatrix} F_{d1} + F_{comp} - F_{f1} \\ -F_{d2} - F_{f2} \\ F_a \end{bmatrix}, \mathbf{u} = \begin{bmatrix} x_{b1} \\ x_{b2} \\ x_s \end{bmatrix}
\end{aligned} \tag{4}$$

where x_s and x_{bi} are respectively the displacements of the stage and bearings.

In an ideal VAN stage, one would expect a perfectly symmetric system such that the mass, damping, stiffness and friction properties of both bearings are exactly the same. In this case, F_{comp} can be eliminated and $F_{d1} = F_{d2} = A \sin(2\pi ft)$ can be used to achieve the purposes of VAN. This ideal situation is evaluated in simulations for a PID controlled VAN stage modeled as shown in Figure 5 with $m_s = 1$ kg, $m_{b1} = m_{b2} = 0.25$ kg, $k_1 = k_2 = 5$ N/ μ m, $c_1 = c_2 = 2 \times 10^{-5}$ N·s/ μ m (corresponding to 1% damping ratio) and $F_{f1} = F_{f2} = F_f$. To allow for a direct comparison with the numerical examples presented in Section 2, the total mass of the VAN stage (i.e., $m = m_s + m_{b1} + m_{b2}$), its P and D controller gains, and its frictional parameters are exactly the same as used in Section 2; the selected stiffness and damping ratios are consistent with those of a properly sized flexure joint (see Section 4). The most conservative integral gain of Figure 2 (i.e., $K_I = 10$ N/(μ m·s)) is used for the VAN stage in all simulations.

Figure 6 shows the settling times of VAN as functions of A and f for the same test cases as in Figure 3 (in response to 50 nm and 500 nm step commands) using various values of the uncertainty parameter, δ . For all four values of δ applied to both step commands, the settling time characteristics can be said to have a cliff, a canyon and a plain. At very low amplitudes and frequencies, the addition of dither produces little or no effect in

reducing settling time (as is predicted by the simple analysis in Appendix A). But as A and f are increased, there is a sudden drop in settling time over the cliff to its lowest levels (i.e., the canyon). As A and f are increased beyond their optimal values, the settling time increases slightly after which it remains largely unchanged (i.e., the plain). The canyon provides optimal settling times at relatively low A and f values. Therefore, operating within the canyon region could also provide benefits with regard to the reduction of heat and wear caused by dither. However, the challenge is that the canyon is very narrow and close to the cliff; settling times could rise sharply if the optimal dithering amplitudes or frequencies are not precise. The plain is excellent in terms of robustness because it provides a wide range of A and f values to choose from without losing performance. Notice also that for the four values of δ , the characteristics of the cliff, canyon and especially the plain are very similar. This means that with sufficiently high A and f values, the performance of VAN is highly insensitive to uncertainties in σ_0 . This fact can be verified based on the theoretical example presented in Appendix A. It shows that, with sufficiently high amplitudes and frequencies of dither, the averaged dynamics does not exhibit any pre-rolling/pre-sliding stiffness behavior and all nonlinear aspects of friction (based on the Dahl model) are linearized. As a result, Problems I and II are eliminated without requiring high integral gain or model-based switching of control gains, and uncertainties in σ_0 do not influence controller performance.

To further illustrate this point, Figure 7 shows the settling responses of VAN to the same step commands of Figure 3 for various values of δ using $A = 3$ N and $f = 500$ Hz (as highlighted by the red dot on each subplot in Figure 6). Compared to the PID controller (with $K_I = 10$ N/($\mu\text{m}\cdot\text{s}$) in Figure 2) and NIASA (in Figure 3), VAN demonstrates superb robustness and settling performance that is very close to the benchmark settling

performance of Figure 2 achieved under zero friction conditions, regardless of frictional uncertainties and magnitudes of reference signals.

Figure 5. A three-mass model of the proposed VAN stage.

Figure 6. Effect of dithering amplitude (A) and frequency (f) on settling time of VAN stage in response to (a) 50 nm, and (b) 500 nm step commands using different values of uncertainty δ in the Dahl friction model. (The red dot represents the operating point of $A = 3$ N and $f = 500$ Hz).

Figure 7. Settling results of VAN in response to (a) 50 nm, and (b) 500 nm step commands using different values of initial contact stiffness in the Dahl model. Note that $A = 3$ N and $f = 500$ Hz. Plots of the benchmark cases from Figure 2 and the nominal cases of NIASA from Figure 3 have been included for the sake of comparison.

3.2.2. Compensation of Parasitic Vibration caused by Non-idealities

In Section 3.2.1, the superb performance and robustness of VAN was demonstrated under the ideal situation where the parameters and friction conditions of both bearings were exactly the same. In reality, non-idealities exist (e.g., different frictional parameters, dithering forces, manufacturing and assembly errors, etc.), which lead to unbalanced reaction forces and parasitic vibration that may adversely affect precise positioning of the stage.

In the preliminary version of this paper [30], for dealing with parasitic vibration, an on-off regulation technique was proposed where dither was applied and then suddenly turned off after duration T , starting when the reference command reached the target position. Significant improvements in settling performance were demonstrated in both simulations and experiments using the on-off regulation approach. However, it was observed that the abrupt switching introduced some undesirable transients which could have large amplitudes if the damping of the stage is small. Moreover, the settling performance was highly dependent on the cutoff time, T , requiring one to identify the optimal T (based on a large number of trials) for each position command, x_r .

To address these issues with the on-off control technique, a control scheme for generating F_{comp} to eliminate parasitic vibration due to non-idealities is proposed, as shown in Figure 8. The signal x_r is the reference position of the table, $F_d = A\sin(2\pi ft)$ is the dithering force which is applied 180° out of phase to both bearings, reflecting the ideal case. A harmonic cancelling (HC) controller [39], is used to generate F_{comp} as a function of the position error of the stage, $e = x_r - x_s$. The generated F_{comp} is then added to the dithering force for bearing 1. Note that a HC controller is a type of repetitive controller [39], designed based on the internal model principle which states that a controller must contain a model of any disturbance that must be rejected with zero steady-state error [40]. In the case of the VAN stage, the disturbance force to be rejected is the un-balanced reaction force due to dither. It can be assumed to be dominated by a harmonic signal at frequency f . Therefore, the transfer function for the HC controller is given by

$$G_{HC}(s) = K_{HC} \frac{s}{s^2 + \omega^2} \quad (5)$$

representing the transfer function of a harmonic signal at frequency f , where s is the Laplace variable, $\omega = 2\pi f$, and K_{HC} is a tunable gain. For any $K_{HC} > 0$, the magnitude of $G_{HC}(s)$ is theoretically infinite at $s = j\omega$ (where j is the unit imaginary number), ensuring that any disturbance (and associated error, e) occurring at frequency f is rejected with zero steady-state error, by modulating the amplitude and phase of F_{d1} relative to F_{d2} .

Figure 8. Block diagram of the proposed control scheme for compensating parasitic vibration caused by non-idealities.

To evaluate the proposed control scheme in the presence of non-idealities, let us consider a scenario where the stiffness and damping coefficients of the two compliant joints differ by 10% due to manufacturing errors; i.e., $k_2 = 1.1 \times k_1 = 5.5 \text{ N}/\mu\text{m}$ and $c_2 = 1.1 \times c_1 = 2.2 \times 10^{-5} \text{ N}\cdot\text{s}/\mu\text{m}$, while the other parameters remain exactly the same as used in Figure 7, with $\delta = 0$. Figure 9 (a) compares the settling behavior of VAN in response to a 50 nm step command with different repetitive controller gains; **very similar results are obtained for the 500 nm case, and are excluded for the sake of brevity**. Note that without the HC controller (i.e., $K_{HC} = 0$), the VAN stage suffers from parasitic vibration with a peak-to-peak amplitude of 40 nm, causing it to not settle within the desired window. By introducing the

HC controller, the parasitic vibration is gradually reduced, allowing the stage to settle. Notice that the rate of attenuation of the parasitic vibration gets faster as K_{HC} is increased. However, K_{HC} cannot be tuned arbitrarily high because a high K_{HC} could introduce instabilities in the HC controller. Therefore, the same type of conservatism that is applied to tuning the integral gain of a PID-type LTI controller must be applied to tuning K_{HC} . Note that the large overshoots experienced when the HC controller is active occur because dither is initiated at the same time as the step command, and the HC controller takes a while to suppress the resulting parasitic vibration. Figure 9 (b) compares the position error of initial step for $K_{HC} = 4 \times 10^3 \text{ N}/\mu\text{m}$ shown in Figure 9 (a) to subsequent steps, applied after the HC controller has taken effect; it shows that the overshoots reduce over time. This means that, to avoid large overshoots, dither could be applied preemptively, allowing resultant parasitic vibration to settle before position commands are issued.

(a)

Figure 9. (a) Settling results of VAN (in the presence of non-idealities) in response to a 50 nm step command with different HC controller gains; (b) position errors of three successive step commands of amplitude 50 nm with $K_{HC} = 4 \times 10^3 \text{ N}/\mu\text{m}$; note that $\delta = 0$, $A = 3 \text{ N}$ and $f = 500 \text{ Hz}$ are used in all simulations.

4. Design and Experimental Evaluation of a Prototype VAN Stage

4.1. Design of a Prototype VAN Stage

A prototype VAN stage is designed according to the general concept described in Section 3.1. Figure 10 shows the CAD drawing of the prototype stage; it has a 2 kg moving mass and 50 mm travel range. The stage is guided by a pair of high-rigidity preloaded linear ball bearings with end seals (THK, SR-15SB), riding on a super-precision grade rail, lubricated using grease (THK, AFB-LF). An air core linear motor (Aerotech, BLMUC-95)

with 162 N peak and 23 N continuous force limits is employed to drive the stage. The table position is measured using a linear encoder system (Renishaw, T1000 read head and RGSZ20 scale) with a post-interpolation resolution of 4.88 nm (notice that the linear encoder is mounted on the side of the table that is not visible in Figure 10). A pair of preloaded piezo stack actuators (PI, P-842.10) is selected to provide the dithering forces; each actuator has a travel range and blocking force of up to 15 μm and 300 N, respectively, with a maximum operating frequency of 6 kHz under unipolar operation (0 – 100 V). Two flexure mechanisms, made of AISI 304 stainless steel, are designed with their central platforms each connected to the stage and their outer platforms connected to their associated linear ball bearings. Their roles are to provide sufficient compliance in the direction of dither (i.e., axial direction) to allow displacements in the order of 10 microns while remaining stiff in the off-axis (i.e., lateral and vertical) directions, such that the high stiffness of the linear ball bearings is not unduly sacrificed. Note that dithering amplitudes in the order of 10 microns are considered to be sufficient because the pre-rolling/pre-sliding regime of rolling bearings becomes most pronounced within 5 – 10 microns of the target position [13].

Table 2 summarizes the stiffness values of each flexure (obtained via finite element analysis using SolidWorks[®]), the stiffness values of each ball bearing (sourced from its manufacturer's catalog), and the combined stiffness values of each flexure and ball bearing (which are in series). Notice that the vertical and lateral stiffness values of each flexure are designed to be much higher than its axial stiffness. Consequently, the combined stiffness of the flexure and linear ball bearing in the vertical and lateral directions have the same order of magnitudes as those of the linear bearing alone. Figure 11 shows the force-displacement

characteristics of the piezo stack actuator with unipolar operation at different voltage levels, obtained from its manufacturer's datasheets [41]. With its axial stiffness of $6.9 \text{ N}/\mu\text{m}$, the flexure provides $11 \mu\text{m}$ of displacement (with 77 N force) at the maximum operating voltage of the piezo actuators (i.e., 100 V). Using SolidWorks[®], the maximum von Mises stress of the flexure (with the maximum applied load of 77 N) is calculated as $S_{max} = 43 \text{ MPa}$. This gives a safety factor of 4.8 with regard to the yield strength S_y of the selected material (i.e., 304 stainless steel), see Table 3. In the presence of cyclic stresses (as is the case with dither), the stress-life method can be employed to determine the strength of materials against fatigue [42]. Infinite fatigue life is guaranteed if the maximum stress is below the endurance limit, S_e , which can be obtained from the Marin equation [43]

$$S_e = p_a p_b p_c p_d p_e p_f S'_e \quad (6)$$

where S'_e is the nominal endurance limit of a carefully prepared rotary-beam specimen under closely controlled conditions; p_a , p_b , p_c , p_d , p_e and p_f are respectively Marin factors that help account for the effects of surface condition, size, loading, reliability, temperature, and miscellaneous items. The nominal endurance limit is estimated as half of the ultimate tensile strength, S_{ut} , as shown in Table 3. The surface factor of a machined structure is given by [42]

$$p_a = 4.51 S_{ut}^{-0.265} \quad (7)$$

while the size factor based on the designed flexure can be calculated using [42]

$$\begin{aligned} p_b &= 1.24 d_e^{-0.107} \\ d_e &= 0.808 \sqrt{h d} \end{aligned} \quad (8)$$

where d_e is the equivalent diameter of the thin flexure hinge (of rectangular section) with height (h) of 19 mm and thickness (d) of 1 mm. A reliability factor of $p_e = 0.814$ is used (based on 99% reliability) and the remaining factors are set equal to 1, assuming standard operating conditions. The estimated Marin factors and strength properties are summarized in Table 3. Note that the endurance limit of 304 stainless steel (i.e., 195.4 MPa) is much higher than the maximum von Mises stress of 47 MPa. Thus, even under the worst-case loading, infinite life against fatigue failure is guaranteed for the designed flexure mechanism.

Figure 10. CAD model of the prototype VAN stage (note: linear encoder is mounted on the other side of the table).

Table 2. Stiffness values of each linear ball bearing, flexure and their combination (i.e., ball bearing + flexure in series) [$\text{N}/\mu\text{m}$].

	Axial	Vertical	Lateral
Ball bearing	N/A	175	58.3
Flexure	6.9	310	331
Combined	6.9	111.9	49.6

Figure 11. Force-displacement relationship of piezo stack actuators at different voltage input levels superimposed with axial force-displacement relationship of flexure mechanism obtained from finite element analysis using SolidWorks[®].

Table 3. Strength properties of the designed flexure.

$S_{max} = 47 \text{ MPa}$	$p_a = 0.86$
$S_{ut} = 517 \text{ MPa}$	$p_b = 1.08$
$S_y = 207 \text{ MPa}$	$p_c = 1$
$S'_e = 258.5 \text{ MPa}$	$p_d = 1$
$S_e = 195.4 \text{ MPa}$	$p_e = 0.814$
$d_e = 3.52 \text{ mm}$	$p_f = 1$

4.2. Experimental Set up and System Identification

Figure 12 shows an in-house-built prototype of the VAN stage described in Section 4.1. The linear motor is driven in force control mode by a PWM amplifier (Aerotech, Soloist CP10). The piezo stack actuator pair is driven separately by voltage amplifiers (PI, E505) which each have an amplification ratio of 10:1.

Figure 13 shows the magnitude and phase plots of frequency response functions (FRFs) measured from the stage; each FRF is obtained by applying constant-amplitude sinusoidal voltage input commands with varying frequencies to the piezo actuator and recording the corresponding output displacement magnitudes of the stage's position using the linear encoder. In both FRFs, a low frequency resonance peak is observed (around 83 Hz), which is characterized as the Dahl resonance of the nonlinear elastic friction in the pre-rolling regime [18]. The higher resonance peaks at around 560 Hz are due to the yaw mode of the stage. Notice that, though the input commands for the two piezo actuators are identical, the obtained FRFs have large discrepancies in terms of both magnitude and phase.

At lower frequencies, the magnitudes differ by a constant offset value but the phase difference is close to the desired 180° . On the other hand, at high frequencies, their magnitudes are closer in value, but their difference in phase is further away from 180° . The mismatched dynamics of the two piezo actuators can be attributed to non-uniformities in the mechanical properties and frictional behaviors of the two linear ball bearings, their flexures and piezo actuators, as discussed in Section 3.2.1.

Figure 12. Assembled prototype VAN stage (note: linear encoder is mounted on the other side of the table).

Figure 13. Measured open loop FRFs from piezo actuator voltage (input) to table position (output). Note that 180° is added to the phase plot of piezo actuator 1 to facilitate easy comparison.

4.3. Experimental Results

The performance of the prototype VAN stage is experimentally evaluated using its settling behavior with and without dither. The control scheme described in Section 3 is implemented on a real-time controller (dSPACE, DS1007) running at 10 kHz sampling frequency. The servo controller of the stage is an LTI PID-type controller with velocity and acceleration feed forward, tuned to have a closed loop bandwidth of 180 Hz using the traditional loop shaping approach, without any information about the applied dithering forces or friction. Note that the servo controller cannot compensate for any un-balanced reaction forces caused by dither with frequencies above its bandwidth of 180 Hz.

A typical 5 mm “move-and-settle” profile shown in Figure 14 is used as the reference command trajectory. It has maximum velocity and acceleration values of 70 mm/s and 1500 mm/s², respectively. The settling portion, after the reference command reaches the target position as indicated by the red dashed line, is considered in this paper.

Figure 15 (a) compares the settling performance of the stage into a ± 100 nm window for three cases: no dither, open loop controlled dither (i.e., without the HC controller) and dither with the proposed control scheme having the HC controller. The settling portion is highlighted in the subplot. Dither input with peak-to-peak amplitude of 20 V (corresponding to 15 N in applied force and approximately 2 μ m in bearing displacement) is used. It is known that in general higher frequency dithering is more effective in mitigating nonlinear effects of friction [23], [28]. Therefore, the dithering frequency is set to 500 Hz, in order to attain high frequency dithering while avoiding the resonance peaks around 560 Hz in the dynamics of the piezo actuators (see Figure 13). Without dither, the stage takes 93.3 ms to settle, which is very long relative to the total duration of the reference command trajectory (i.e., 145 ms). With open loop controlled dither, as expected, the stage experiences parasitic vibration, especially in the settling region, with amplitude of about 300 nm (causing the stage to not settle into the desired window). Figure 15 (b) shows the frequency spectrum of the in-position vibration amplitude using the position error signals between 200 and 300 ms. The large peak (with amplitude of 270 nm) occurring at 500 Hz indicates that the parasitic vibration is mainly concentrated at the dithering frequency. Note that the frequency spectrum of the case without dither has a very large DC value (compared to the case with dither) because the stage is still moving slowly to the target position due to the effects of pre-rolling/pre-sliding

friction. The proposed control scheme with the HC controller is then applied, helping the stage with dither to settle within 41.6 ms, which is 55.4% faster than the case without dither. Notice that, thanks to the HC controller, the dominant peak at 500 Hz is eliminated, and that the overall motion trajectory of the stage is largely unaffected when dither is applied (as the peak and RMS errors remain almost the same). However, compared to the case without dither, the stage with dither moves faster to the target position near the end of the motion command (where the reference velocities are small and pre-rolling/pre-sliding friction dominates). Dither effectively mitigates the highly nonlinear friction that dominates once the stage moves into the pre-rolling/pre-sliding regime, leading to smaller position errors at the start of the settling region as shown in Figure 15 (a). This indicates that the proposed VAN stage has the potential of improving not only point-to-point positioning but also tracking performance.

Figure 16 plots the settling performance of the prototype stage with dither (combined with the HC control scheme), in comparison with the case without dither, based on 50 trials at random positions of the table within its travel range. It is observed that the settling performance of the stage varies, due to differences in friction dynamics from position to position along the travel range of the stage. Without dither, the mean settling time into the ± 100 nm window is 82.5 ms; with dither, the stage is able to achieve a mean settling time of 39.6 ms, i.e., a 52% reduction in mean settling time. Note that though the settling time with dither is less than 50 ms for most ($\sim 85\%$) of the trials, it increases significantly on occasion. With the HC controller, the dominant peak is suppressed; however, some low frequency disturbances show up when dither is applied, possibly due to hysteresis in the piezo actuators. Since the FB controller of the linear motor is tuned

independently of the applied dither, it amplifies some of these disturbances which show up as peaks between 50 Hz and 200 Hz (see Figure 15 (b)). Occasionally, the amplitudes of these peaks become large enough (compared to the desired settling window) such that the settling performance of the stage is affected. Figure 17 shows a sample case where the settling time of the stage with dither is large. Using a strict metric of settling, the settling time is 79.4 ms, although the stage momentarily steps out of the desired window for only 1 ms. Therefore, for practical purposes the stage could be said to have settled after it first enters the ± 100 nm window at 44.3 ms.

Figure 14. Commanded position and velocity of a typical 5 mm “move-and-settle” motion.

The red dashed line indicates the start of the settling portion.

(a)

(b)

Figure 15. (a) Typical responses of no dither, dither without and dither with the proposed HC control scheme (the red box shows the settling portion); (b) frequency spectrum of position error signals (from 200 to 300 ms).

Figure 16. Comparison of settling times into the ± 100 nm window for the case with and without dither based on 50 trials at random positions of the table within its travel range.

Note that the HC controller is used with dither in all cases

Figure 17. Typical response of the stage with dither (and the HC controller) where large settling times are observed; only the settling portion of the position error signals is extracted

4.4. Remarks on the Effects of Dither on Heat and Wear

As was mentioned in Section 3.1, one worry about dither is its potential to generate heat and cause wear of mechanical components. So it is important to investigate the significance of these concerns for VAN. The power consumption (P_{PZT}) of piezo stack actuators in sinusoidal operation can be approximated as [41]

$$P_{PZT} = f \cdot C_{PZT} \cdot V_{pp}^2 \quad (9)$$

where C_{PZT} is the capacitance of the piezo actuator, f is the operating frequency and V_{pp} is the applied (peak-to-peak) voltage. The total power consumption of the two piezo stacks of the prototype VAN stage for the 5 mm move-and-settle motion can be calculated as 0.6 W (i.e., 0.3 W for each actuator). It is suggested that up to 10% of the electrical power is converted into heat [41], which is less than 0.06 W. In other words, only about 0.015 J of

thermal energy is added to the stage for each 5 mm move-and-settle motion (based on 145 ms for the reference trajectory and an additional 100 ms for settling). Figure 18 shows the temperature of the table and piezo actuators, measured using a thermocouple (Omega, type-K), for continuous dithering (at $V_{pp} = 20$ V and $f = 500$ Hz) over 30 minutes, starting from thermal equilibrium at room temperature. It can be seen that both piezo actuators heat up by about 0.2 °C, but the table temperature remains reasonably constant. This shows that the thermal impact of continuous dithering is negligible for the prototype stage. It is also worth mentioning that continuous dithering (e.g., for 30 minutes) is not necessary in many situations, like point-to-point positioning.

In order to evaluate the potential bearing wear caused by continuous dithering, the vibration amplitude of the linear ball bearing, which is approximately 2 μm based on the applied dithering amplitude of 20 V (see Figure 11), is used. The additional travel of the bearing caused by dither for each 5 mm move-and-settle motion is calculated as 0.49 mm. This corresponds to 9% reduction of the bearing's service life, compared to the case without dither, as calculated in Appendix B using equations provided by the bearing manufacturer [44]. Note that the nominal life of the linear ball bearing is more than 74,000 km, corresponding to over 14.8 billion 5 mm move-and-settle motions (or over 115 years continuous operation). Therefore, a 9% reduction in service life is a moot point in practice, especially considering the large increases in productivity that can be gained by reducing settling times using VAN.

Figure 18. Temperature of the table and piezo actuators of the prototype VAN stage during continuous dithering at 20 V and 500 Hz for 30 minutes

5. Conclusions and Future Work

In this paper, a novel approach for mitigating the pre-rolling/pre-sliding-friction-induced slow settling problem of rolling bearing nanopositioning stages using dither has been presented. The proposed method, called vibration assisted nanopositioning (VAN), synergistically combines electromechanical design and control to allow dithering forces to be applied to a precision stage without causing excessive parasitic vibration, heat or wear. Specifically, it proposes the use of two short-stroke actuators (e.g., piezo or voice coil actuators), acting opposite to each other, to apply dithering forces directly to the rolling bearings and hence minimize unbalanced reaction forces due to dithering; the rolling bearings are connected to the stage using compliant joints (e.g., flexure joints) to facilitate dithering. Moreover, it proposes a method for using a harmonic cancelling controller to

compensate any residual parasitic vibration, caused by non-idealities in the mechanical design of the stage.

Simulations have been used to demonstrate VAN's potential for superior performance and robustness compared to a recently-proposed model-based friction compensating method. A prototype VAN stage has been designed and built based on the proposed concept. Preliminary experiments conducted using the designed prototype have demonstrated up to 52% reduction in mean settling time in point-to-point positioning based on 50 random trials. The additional heat and potential wear caused by dithering have been shown to be insignificant in practice. Future work will investigate alternative control techniques for suppressing parasitic vibration without causing the slight increases in low frequency vibration observed in experiments. Moreover, the influence of dithering parameters (e.g., frequency and amplitude) on settling time will be explored both theoretically and experimentally.

Acknowledgements

The authors are grateful to Professor Shorya Awtar for providing valuable feedback on the prototype design and to Aerotech Inc. for donating some of the key components used in building the prototype stage. This work is funded by National Science Foundation's Award CMMI #1562297: Vibration Assisted Nanopositioning: An Enabler of Low-cost, High-throughput Nanotech Processes.